



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 94276

TO: Manjunath N Rao  
Location: CM1/10D01/10A11  
Art Unit: 1652  
Wednesday, May 28, 2003

Case Serial Number: 040863

From: Susan Hanley  
Location: Biotech-Chem Library  
CM1-6B05  
Phone: 305-4053

susan.hanley@uspto.gov

### Search Notes

10/040863

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:01:52 ; Search time 37.6369 seconds  
(without alignments)  
1932.533 Million cell updates/sec

Title: US-10-040-863-10

Perfect score: 1896

Sequence: 1 LQQRIVKQLPSKELPMTT.....RSHFLKAKGVTCYVAGRAF 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1896	100.0	353	11 055025	O55025 rattus norv
2	1896	100.0	380	11 09JK44	Q9JK44 rattus norv
3	1747	92.1	354	11 035087	Q35087 rattus norv
4	1747	92.1	354	11 09R275	Q9R275 rattus norv
5	1538.5	81.1	347	11 09JL27	Q9JL27 mus musculu
6	1534.5	80.9	347	11 070504	O70504 mus musculu
7	1483.5	78.2	321	11 0920W2	Q920W2 mus spicile
8	1483.5	78.2	321	11 091VF0	Q91VF0 mus musculu
9	1468.5	77.5	321	11 0920W3	Q920W3 mus musculu
10	1400.5	73.9	343	6 077486	O77486 gorilla gor
11	1397.5	73.7	338	6 09TUD5	Q9TUD5 pan troglod
12	1397.5	73.7	343	6 077485	O77485 pan troglod
13	1396.5	73.7	343	6 09TUD3	Q9TUD3 gorilla gor
14	1391.5	73.4	347	6 029505	Q29505 oryctolagus
15	1389.5	73.3	343	6 077487	O77487 pongo pygma
16	1382.5	72.9	332	4 099450	Q99450 homo sapien

17	1382.5	72.9	343	6 077712	O77712 cercopithec
18	1379.5	72.8	343	6 09TTC7	Q9TTC7 hylobates i
19	1346	71.0	368	11 0920T3	Q920T3 mus musculu
20	1343	70.8	299	11 09WCE6	Q9WCE6 rattus norv
21	1342	70.8	368	11 0920T2	Q920T2 mus spicile
22	1339	70.6	346	4 014338	Q14338 homo sapien
23	1338	70.6	368	11 0920T4	Q920T4 mus musculu
24	1330	70.1	368	11 091V73	Q91V73 mus musculu
25	1319	69.6	273	6 09TSL5	O9TSL5 macaca fasc
26	1307	68.9	273	6 09TTC9	Q9TTC9 eulemur ful
27	1262	66.6	273	6 09TTD0	Q9TTD0 callithrix
28	1229	64.8	348	6 09TUD1	Q9TUD1 macaca mula
29	1214	64.0	368	6 09TTY3	Q9TTY3 bos taurus
30	1207	63.7	332	6 09N266	Q9N266 macaca fasc
31	1109	58.5	377	11 09JL28	Q9JL28 mus musculu
32	1109	58.5	377	11 09T327	P9T327 mus musculu
33	1099	58.0	377	11 0920W5	Q920W5 mus musculu
34	1099	58.0	377	11 0920W4	Q920W4 mus spicile
35	1086.5	57.3	360	6 09TTY7	Q9TTY7 bos taurus
36	1051.5	55.5	366	6 09TUD4	Q9TUD4 gorilla gor
37	1050.5	55.4	365	6 09TUE6	Q9TUE6 eulemur ful
38	1048.5	55.3	365	6 09TUE7	Q9TUE7 hylobates i
39	1048	55.3	366	6 09TUD2	Q9TUD2 macaca mula
40	1047.5	55.2	353	6 09N270	Q9N270 pongo pygma
41	1047.5	55.2	365	4 08WXB6	Q8WXB6 homo sapien
42	1047	55.2	365	6 09N267	Q9N267 macaca fasc
43	1046.5	55.2	365	4 014505	O14505 homo sapien
44	1043.5	55.0	366	6 09TUD6	Q9TUD6 pan troglod
45	1042.5	55.0	365	4 014506	O14506 homo sapien

ALIGNMENTS

RESULT 1

ID: O55025 PRELIMINARY: PRT: 353 AA.

AC O55025;  
DT 01-JUN-1998 (TRENBLrei. 06, Created)  
DT 01-JUN-1998 (TRENBLrei. 06, Last sequence update)  
DE Alpha 1,2 fucosyltransferase (Fragment).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BUFFALO RAT;  
RX MEDLINE=98342056; PubMed=9675030;  
RA Sherwood A.L., Holmes E.H.;  
RT "Cloning and expression of the catalytic domain from rat hepatoma H35  
cell GDP-fucose: GM1 alpha 1-->2fucosyltransferase, an enzyme which is  
activated during early stages of chemical carcinogenesis in rat  
liver".  
RL Arch. Biochem. Biophys. 355:215-221(1998).  
DR EMBL: AF042743; AAC14695.1;  
DR InterPro: IPR002516; GT11.  
DR Pfam: PF01531; Glyco\_transf\_11;  
KW Glycosyltransferase; Transferase.  
FT NON\_TER 1  
SQ SEQUENCE 353 AA; 39738 MW; 609D64EB22C9585 CRC64;

Query Match 100.0%; Score 1896; DB 11; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4.3e-164;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQQRIVKQLPSKELPMTTMSNGTSPENRRDSEQHNGELRCMFTINSIGRLGNQM 60  
DB 1 LQQRIVKQLPSKELPMTTMSNGTSPENRRDSEQHNGELRCMFTINSIGRLGNQM 60  
QY 61 GYATLAFALRNGLRIATPASHNALAFIFRISLPVLSHTAKKIPWQNYHLNDWMEER 120  
|||||

```
Db 61 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEER 120
QY 121 YRHIPGHEVFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 180
Db 121 YRHIPGHEVFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 180
QY 181 GYHVRGDIYHVMNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCRENINAS 240
Db 181 GYHVRGDIYHVMNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCRENINAS 240
QY 241 RGVVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLAGGDTIYLANITLPDSPF 300
Db 241 RGVVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLAGGDTIYLANITLPDSPF 300
QY 301 LKVFPEAAFLPEWVGIPADLSPLLKALIPACPRSHFHLKAGVTCYVAGRAF 353
Db 301 LKVFPEAAFLPEWVGIPADLSPLLKALIPACPRSHFHLKAGVTCYVAGRAF 353
RESULT 2
Q5UK44 PRELIMINARY; PRT: 380 AA.
ID Q5UK44
AC Q5UK44;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FISCHER;
RA Sherwood A.L.; Stroud M.R.; Levery S.B.; Holmes E.H.;
RT "An amino acid region at the N-terminus of rat hepatoma H35 cell GDP-
RT fucose:GM1 alpha 1-2 fucosyltransferase is required for optimum enzyme
RT activity and interaction with lipids."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF264005; AAF7200.1; -.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 380 AA; 42742 MW; 920692C96689C05C CRC64;

Query Match 100.0%; Score 1896; DB 11; Length 380;
Best Local Similarity 100.0%; Pred. No. 4.7e-164;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQORIVKLOPLSEKELPMTQMSSGNTESPMRRDSEQHNGELRGMTINSIGRLGNQM 60
Db 28 LQORIVKLOPLSEKELPMTQMSSGNTESPMRRDSEQHNGELRGMTINSIGRLGNQM 87
QY 61 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEER 120
Db 88 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEER 147
QY 121 YRHIPGHEVFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 180
Db 148 YRHIPGHEVFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 207
QY 181 GYHVRGDIYHVMNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCRENINAS 240
Db 208 GYHVRGDIYHVMNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCRENINAS 267
QY 241 RGVVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLAGGDTIYLANITLPDSPF 300
Db 268 RGVVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLAGGDTIYLANITLPDSPF 327
QY 301 LKVFPEAAFLPEWVGIPADLSPLLK 326
Db 328 LKVFPEAAFLPEWVGIPADLSPLLK 353

RESULT 4
Q9R275 PRELIMINARY; PRT: 354 AA.
ID Q9R275
AC Q9R275;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase B.
GN FTB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BDIX;
RX MEDLINE=21099374; PubMed=11179967;
```

## RESULT 3

```
Q55087 PRELIMINARY; PRT: 354 AA.
ID Q55087
AC Q55087;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase.
GN FTB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON CANCER;
RA Soejima M.; Wang B.; Koda Y.; Kimura H.;
RT "Two distinct rat GDP-L-fucose:b-D-galactoside 2-a-L-
RT fucosyltransferase genes."
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB006138; BAA21742.1; -.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 354 AA; 39995 MW; 8636444888215BA1 CRC64;

Query Match 92.1%; Score 1747; DB 11; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.5e-150;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQORIVKLOPLSEKELPMTQMSSGNTESPMRRDSEQHNGELRGMTINSIGRLGNQM 60
Db 28 LQORIVKLOPLSEKELPMTQMSSGNTESPMRRDSEQHNGELRGMTINSIGRLGNQM 87
QY 61 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEER 120
Db 88 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEER 147
QY 121 YRHIPGHEVFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 180
Db 148 YRHIPGHEVFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 207
QY 181 GYHVRGDIYHVMNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCRENINAS 240
Db 208 GYHVRGDIYHVMNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCRENINAS 267
QY 241 RGVVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLAGGDTIYLANITLPDSPF 300
Db 268 RGVVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLAGGDTIYLANITLPDSPF 327
QY 301 LKVFPEAAFLPEWVGIPADLSPLLK 326
Db 328 LKVFPEAAFLPEWVGIPADLSPLLK 353

RESULT 4
Q9R275 PRELIMINARY; PRT: 354 AA.
ID Q9R275
AC Q9R275;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase B.
GN FTB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BDIX;
RX MEDLINE=21099374; PubMed=11179967;
```

RA Bureau V., Marionneau S., Caillaud-Thomas A., Le Moullac-Vaidye B.,  
RA Liehr T., Le Pendu J.,  
RT "Comparison of the three rat GDP-L-fucose: beta-D-galactoside 2-  
RL Eur. J. Biochem. 268:1006-1019(2001).  
DR EMBL: AF131238; AAD24469.1; -.  
DR InterPro: IPR002516; GT\_11.  
DR Pfam: PF01531; Glyco\_transf\_11; 1.  
KW Glycosyltransferase; Transferrase  
SQ SEQUENCE 354 AA; 39983 MW; 123E8C8379E8559E CRC64;  
  
Query Match 92.1%; Score 1747; DB 11; Length 354;  
Best Local Similarity 100.0%; Pred. No. 1.5e-150;  
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 LQORIVKLOPLSEKELPMTQMSSGNTSEPMRRDSEHQNGELRGMTINSIGRLGNQM 60  
DB 28 LQORIVKLOPLSEKELPMTQMSSGNTSEPMRRDSEHQNGELRGMTINSIGRLGNQM 87  
  
QY 61 GEYATLALARMNGRLAFIPASMHNALAFIRISLPVLSHSDTAKKIPWQNYHLNDWMEER 120  
DB 88 GEYATLALARMNGRLAFIPASMHNALAFIRISLPVLSHSDTAKKIPWQNYHLNDWMEER 147  
  
QY 121 YRHIPGHFVRFYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 180  
DB 148 YRHIPGHFVRFYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 207  
  
QY 181 GVHVRGDIYVHVPNVKGVADRGYLEKALDMFRARYSPVFWVTSNGMAWCRENINAS 240  
DB 208 GVHVRGDIYVHVPNVKGVADRGYLEKALDMFRARYSPVFWVTSNGMAWCRENINAS 267  
  
QY 241 RGDVFPAGNGIEGSPAKDFALLTQCNTMTICTGFWAAYLAGGDTIYLYANTLPDPSF 300  
DB 268 RGDVFPAGNGIEGSPAKDFALLTQCNTMTICTGFWAAYLAGGDTIYLYANTLPDPSF 327  
  
QY 301 LKVFKEPAALFPEWVGIPADLSPLK 326  
DB 328 LKVFKEPAALFPEWVGIPADLSPLK 353  
  
RESULT 5  
Q9JL27 PRELIMINARY; PRT; 347 AA.  
ID AC Q9JL27  
DC Q9JL27  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)  
DE Alpha(1,2)fucosyltransferase FUT2 (EC 2.4.1.69).  
GN FUT2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/OLA;  
RX MEDLINE=21316545; PubMed=11323419;  
RA Domino S.E., Zhang L., Lowe J.B.;  
RT "Molecular Cloning, Genomic Mapping, and Expression of Two Secretor  
RT Blood Group alpha (1,2)fucosyltransferase Genes Differentially  
RT Regulated in Mouse Uterine Epithelium and Gastrointestinal Tract.";  
RL J. Biol. Chem. 276:23748-23756(2001).  
DR EMBL: AF214656; AAF45146.1; -.  
DR MGD; MGI:109374; Fut2.  
DR InterPro: IPR002516; GT\_11.  
DR Pfam: PF01531; Glyco\_transf\_11; 1.  
KW Glycosyltransferase; Transferrase.  
SQ SEQUENCE 347 AA; 39243 MW; DD5F441046C8E79 CRC64;  
  
Query Match 81.1%; Score 1538.5; DB 11; Length 347;  
Best Local Similarity 89.0%; Pred. No. 1.4e-131;  
Matches 290; Conservative 12; Mismatches 17; Indels 7; Gaps 2;

QY 1 LQORIVKLOPLSEKELPMTQMSSGNTSEPMRRDSEHQNGELRGMTINSIGRLGNQM 60  
DB 28 LQORIVKLOPLSEKELPMTQMSSGNTSEPMRRDSEHQNGELRGMTINSIGRLGNQM 80  
  
QY 61 GEYATLALARMNGRLAFIPASMHNALAFIRISLPVLSHSDTAKKIPWQNYHLNDWMEER 120  
DB 81 GEYATLALARMNGRLAFIPESMHNALAFIRISLPVLSHSDTARRIPWQNYHLNDWMEER 140  
  
QY 121 YRHIPGHFVRFYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 180  
DB 141 YRHIPGHFVRFYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 200  
  
QY 181 GVHVRGDIYVHVPNVKGVADRGYLEKALDMFRARYSPVFWVTSNGMAWCRENINAS 240  
DB 201 GVHVRGDIYVHVPNVKGVADRGYLEKALDMFRARYSPVFWVTSNGMAWCRENINAS 260  
  
QY 241 RGDVFPAGNGIEGSPAKDFALLTQCNTMTICTGFWAAYLAGGDTIYLYANTLPDPSF 300  
DB 261 RGDVFPAGNGIEGSPAKDFALLTQCNTMTICTGFWAAYLAGGDTIYLYANTLPDPSF 320  
  
QY 301 LKVFKEPAALFPEWVGIPADLSPLK 326  
DB 321 LKVFKEPAALFPEWVGIPADLSPLK 346  
  
RESULT 6  
Q70504 PRELIMINARY; PRT; 347 AA.  
ID AC Q70504  
DC Q70504  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)  
DE GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69).  
GN FUT2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ICR; TISSUE=GASTROINTESTINAL TRACT;  
RX MEDLINE=20471982; PubMed=11018479;  
RA Lin B., Hayashi Y., Saito M., Sakakibara Y., Yanagisawa M.,  
RA Iwamori M.;  
RT "GDP-fucose: beta-galactoside alpha1,2-fucosyltransferase, MFUT-II,  
RT and not MFUT-I or -III, is induced in a restricted region of the  
RT digestive tract of germ-free mice by host-microbe interactions and  
RT cycloheximide.";  
RL Biochim. Biophys. Acta 1487:275-285(2000).  
DR EMBL: AF064792; AAC16887.1; -.  
DR MGD; MGI:109374; Fut2.  
DR InterPro: IPR002516; GT\_11.  
DR Pfam: PF01531; Glyco\_transf\_11; 1.  
KW Glycosyltransferase; Transferrase.  
SQ SEQUENCE 347 AA; 39215 MW; 616CCB11581C4179 CRC64;  
  
Query Match 80.9%; Score 1534.5; DB 11; Length 347;  
Best Local Similarity 88.7%; Pred. No. 3.1e-131;  
Matches 289; Conservative 13; Mismatches 17; Indels 7; Gaps 2;



```

QY 181 GVHVRGDIYHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCARENINAS 240
Db 201 GVHVRGDIYHVMPKVKGVVADRGYLEKALDRFARYSSPVFVVTSGMAWCARENINTS 260
QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTMTGTCGIIWAAYLAGGDTIYLANITLPDSPF 300
Db 261 LGDVVFAGNGIEGSPAKDFALLTQCNTHTMTGTCGIIWAAYLAGGDTIYLANITLPDSPF 320
QY 301 LKVFKEPAAFLEPWGIPADLSPLLK 326
Db 321 LKIFKPAFAFLPEWGPADLSPLLK 346

RESULT 7
Q920W2 PRELIMINARY; PRT; 321 AA.
AC Q920W2
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase (EC 2.4.1.69) (Fragment).
GN FUT2.
OS Mus spicilegus (Steppe mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZBN;
RA Liu Y., Kitano T., Shiroishi T., Moriwaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039123; BAB68647.1; -
DR InterPro; IPR002516; GT_11; -
DR Pfam; PF01531; Glyco_transferase.
KW Glycosyltransferase; transferase.
FT NON_TER 1 1
FT NON_TER 321 321
SQ SEQUENCE 321 AA; 36560 MW; 5B7D19BA6D4B5394 CRC64;

```

```

Query Match 78.2%; Score 1483.5; DB 11; Length 321;
Best Local Similarity 88.9%; Pred. No. 1.2e-126;
Matches 280; Conservative 11; Mismatches 17; Indels 7; Gaps 2;

QY 1 LQQRIVKLOPLSEKELPMTTQMSGNTSPERMRRDSEQHNGELRGMTINSIGRLGNQM 60
Db 14 LQQRIVKLOPLSEKEL-QAQMSSLNARTDMQ-----SAKLOGIITINSIGRLGNQM 66

QY 61 GEYATLALAMNGRLATIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEER 120
Db 67 GEYATLALAMNGRLATIPESMHNALAPIFRISLPVLHSDTARRIPQNYHLNDWMEER 126

QY 121 YRHIPGHEVRFYPCSWTFYHHLRPEILKEFTLHDHVEEAQAFLRLGVNGSOPSTFV 180
Db 127 YRHIPGQVRFYPCSWTFYHHLRPEILKEFTLHDHVEEAQAFLRLGVNGSOPSTFV 186

QY 181 GVHVRGDIYHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCARENINAS 240
Db 187 GVHVRGDIYHVMPKVKGVVADRGYLEKALDRFARYSSPVFVVTSGMAWCARENINTS 246

QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTMTGTCGIIWAAYLAGGDTIYLANITLPDSPF 300
Db 247 LGDVVFAGNGIEGSPAKDFALLTQCNTHTMTGTCGIIWAAYLAGGDTIYLANITLPDSPF 306

QY 301 LKVFKEPAAFLEPW 315
Db 307 LKIFKPAFAFLPEW 321

RESULT 8
Q91VF0 PRELIMINARY; PRT; 321 AA.
AC Q91VF0
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase (EC 2.4.1.69) (Fragment).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VARIOUS STRAINS;
RA Liu Y., Kitano T., Koida T., Shiroishi T., Moriwaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039114; BAB68638.1; -
DR EMBL; AB039115; BAB68639.1; -
DR EMBL; AB039116; BAB68640.1; -
DR EMBL; AB039117; BAB68641.1; -
DR EMBL; AB039118; BAB68642.1; -
DR EMBL; AB039119; BAB68643.1; -
DR EMBL; AB039120; BAB68644.1; -
DR EMBL; AB039121; BAB68645.1; -
DR InterPro; IPR002516; GT_11; -
DR Pfam; PF01531; Glyco_transferase.
KW Glycosyltransferase; transferase.
FT NON_TER 1 1
FT NON_TER 321 321
SQ SEQUENCE 321 AA; 36516 MW; 0622D3BB503B72D1 CRC64;

```

```

Query Match 78.2%; Score 1483.5; DB 11; Length 321;
Best Local Similarity 88.6%; Pred. No. 1.2e-126;
Matches 279; Conservative 12; Mismatches 17; Indels 7; Gaps 2;

QY 1 LQQRIVKLOPLSEKELPMTTQMSGNTSPERMRRDSEQHNGELRGMTINSIGRLGNQM 60
Db 14 LQQRIVKLOPLSEKEL-QAQMSSLNARTDMQ-----SAKLOGIITINSIGRLGNQM 66

QY 61 GEYATLALAMNGRLATIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEER 120
Db 67 GEYATLALAMNGRLATIPESMHNALAPIFRISLPVLHSDTARRIPQNYHLNDWMEER 126

QY 121 YRHIPGHEVRFYPCSWTFYHHLRPEILKEFTLHDHVEEAQAFLRLGVNGSOPSTFV 180
Db 127 YRHIPGQVRFYPCSWTFYHHLRPEILKEFTLHDHVEEAQAFLRLGVNGSOPSTFV 186

QY 181 GVHVRGDIYHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCARENINAS 240
Db 187 GVHVRGDIYHVMPKVKGVVADRGYLEKALDRFARYSSPVFVVTSGMAWCARENINTS 246

QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTMTGTCGIIWAAYLAGGDTIYLANITLPDSPF 300
Db 247 LGDVVFAGNGIEGSPAKDFALLTQCNTHTMTGTCGIIWAAYLAGGDTIYLANITLPDSPF 306

QY 301 LKVFKEPAAFLEPW 315
Db 307 LKIFKPAFAFLPEW 321

RESULT 9
Q920W3 PRELIMINARY; PRT; 321 AA.
AC Q920W3
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase (EC 2.4.1.69) (Fragment).
GN FUT2.

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
OX NCBI_TaxId=9598;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=CH16;
RC MEDLINE=2018794; PubMed=10723735;
RX Apoll P.-A., Roubinet F., Desplau S., Mollicone R., Oriol R.,
RA Blancher A.;
RT "Evolution of alpha2-Fucosyltransferase genes in primates: relation
RT between an intronic Alu-Y element and red cell expression of ABH
RT antigens";
RL Mol. Biol. Evol. 17:337-351(2000).
DR EMBL: AF080604; AAF14066.1; -
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Transferase.
SQ SEQUENCE 338 AA; 38428 MW; E7FBA0FFB9C95F9 CRC64;

Query Match      73.7%; Score 1397.5; DB 6; Length 338;
Best Local Similarity 79.4%; Pred. No. 8.7e-119;
Matches 259; Conservative 27; Mismatches 29; Indels 11; Gaps

Qy 1 LQQRIVKLPQLSEKELPMVTQMSSGNSTESPEMRDSDQHGHNGELRGMTFTNSTGRIGNQM 60
   :|||:| |: | | | : | : | : |||:|||||
Db 23 VQQRLLAIQAM--WELPV-----QIPVLASTSKALGPSOLRGMWTNIAIGRLNQM 71

```

```

Qy 61 GEYATLALARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
Db 72 GEYATLYALAKMNGRPAPFPAQMHSHTLAPFRITLPLVLSATASRIPWQNYHLNDWMEEE 131
Qy 121 YRHIPGHVVRTGTPCSTWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSPSTFV 180
Db 132 YRHIPGEYVVRTGTPCSTWTFYHHLRQELQELQFTLHDHVREAAQAFRLGLRVNGSPSTFV 191
Qy 181 GVHVRGDIYVHMPNWKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCNENINAS 240
Db 192 GVHVRGDIYVHMPNWKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCNENIDTS 251
Qy 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGGIWAAYLAGDGTIYLANYLTPDSPF 300
Db 252 HGDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTGGIWAAYLTGDDTIYLANYLTPDSPF 311
Qy 301 LKVPKPEAAFLPEWVGIPADLSPLK 326
Db 312 LKVPKPEAAFLPENMGIAADLSPLK 337

RESULT 12
Q77485 PRELIMINARY: PRT: 343 AA.
ID O77485
AC O77485
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Alpha(1,2)fucosyltransferase.
GN FUT2.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;
RT "The old origin of a null allele se428 of the human ABO-secretor type
RL alpha(1,2) fucosyltransferase gene (FUT2).";
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human secretor blood
RT group alpha(1,2) fucosyltransferase gene (FUT2).";
RL J. Biol. Chem. 270:4640-4649(1995).
DR EMBL; AB015634; BAA31127.1; -.
DR InterPro; IPR002516; GP_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Glyco-transferase.
SQ SEQUENCE 343 AA; 38999 MW; B849D7E152852081 CRC64;

Query Match 73.7%; Score 1397.5; DB 6; Length 343;
Best Local Similarity 79.4%; Pred. No. 8.9e-119;
Matches 259; Conservative 27; Mismatches 29; Indels 11; Gaps 2;

Qy 1 LQORIVKLQPLSEKELPMTTOMSSGNTESPEMRDSEOHNGELRGMTINSIGRLGNQM 60
Db 28 VQORLAKIQAM-WELPV-----QIPVLASTSKALGPSQLRGMTINAIKRLGNQM 76
Qy 61 GEYATLALARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
Db 77 GEYATLYALAKMNGRPAPFPAQMHSHTLAPFRITLPLVLSATASRIPWQNYHLNDWMEEE 136
Qy 121 YRHIPGHVVRTGTPCSTWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSPSTFV 180
Db 132 YRHIPGEYVVRTGTPCSTWTFYHHLRQELQELQFTLHDHVREAAQAFRLGLRVNGSPSTFV 196
Qy 181 GVHVRGDIYVHMPNWKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCNENINAS 240
Db 192 GVHVRGDIYVHMPNWKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCNENIDTS 256
Qy 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGGIWAAYLAGDGTIYLANYLTPDSPF 300
Db 252 HGDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTGGIWAAYLTGDDTIYLANYLTPDSPF 316
Qy 301 LKVPKPEAAFLPEWVGIPADLSPLK 326
Db 312 LKVPKPEAAFLPENMGIAADLSPLK 342

RESULT 14
Q29505 PRELIMINARY: PRT: 347 AA.
ID Q29505
AC Q29505
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)

```

```

Qy 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGGIWAAYLAGDGTIYLANYLTPDSPF 300
Db 257 HGDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTGGIWAAYLTGDDTIYLANYLTPDSPF 316
Qy 301 LKVPKPEAAFLPEWVGIPADLSPLK 326
Db 317 LKVPKPEAAFLPENMGIAADLSPLK 342

RESULT 13
Q9TUD3 PRELIMINARY: PRT: 343 AA.
ID Q9TUD3
AC Q9TUD3
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Alpha (1,2) fucosyl transferase.
GN FUT2.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=ALEXIS;
RX MEDLINE=20188794; PubMed=10723735;
RA Apoll P.-A., Roubinet F., Despiau S., Mollicone R., Oriol R.,
RA Blancher A.;
RT "Evolution of alpha2-Fucosyltransferase genes in primates: relation
RT between an intronic Alu-Y element and red cell expression of ABH
RT antigens.";
RL Mol. Biol. Evol. 17:337-351(2000).
DR EMBL; AF080606; AAF14068.1; -.
DR InterPro; IPR002516; GP_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Transferase.
SQ SEQUENCE 343 AA; 39001 MW; A753375D47AB8C8 CRC64;

Query Match 73.7%; Score 1396.5; DB 6; Length 343;
Best Local Similarity 79.4%; Pred. No. 1.1e-118;
Matches 259; Conservative 26; Mismatches 30; Indels 11; Gaps 2;

Qy 1 LQORIVKLQPLSEKELPMTTOMSSGNTESPEMRDSEOHNGELRGMTINSIGRLGNQM 60
Db 28 VQORLAKIQAM-WELPV-----QIPVLASTSKALGPSQLRGMTINAIKRLGNQM 76
Qy 61 GEYATLALARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
Db 77 GEYATLYALAKMNGRPAPFPAQMHSHTLAPFRITLPLVLSATASRIPWQNYHLNDWMEEE 136
Qy 121 YRHIPGHVVRTGTPCSTWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSPSTFV 180
Db 137 YRHIPGEYVVRTGTPCSTWTFYHHLRQELQELQFTLHDHVREAAQAFRLGLRVNGSPSTFV 196
Qy 181 GVHVRGDIYVHMPNWKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCNENINAS 240
Db 197 GVHVRGDIYVHMPNWKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCNENIDTS 256
Qy 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGGIWAAYLAGDGTIYLANYLTPDSPF 300
Db 257 HGDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTGGIWAAYLTGDDTIYLANYLTPDSPF 316
Qy 301 LKVPKPEAAFLPEWVGIPADLSPLK 326
Db 317 LKVPKPEAAFLPENMGIAADLSPLK 342

RESULT 14
Q29505 PRELIMINARY: PRT: 347 AA.
ID Q29505
AC Q29505
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)

```

01-MAR-2002 (TREMELREL. 20, Last annotation update)  
 Galactoside 2-L-fucosyltransferase 3 (EC 2.4.1.69) (Secretor blood  
 group ALPHA-2-fucosyltransferase) (GDP-L-fucose-beta-D-galactoside  
 2-ALPHA-L-fucosyltransferase) (ALPHA(1,2)FUT) (Fucosyltransferase 3).  
 OS Eryctolagus cuniculus (Rabbit).  
 OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=GASTROINTESTINAL TRACT;  
 RX MEDLINE=96279281; PubMed=8663168;  
 RA Hitoshi S., Kojima K., Kanazawa I., Tsuji S.;  
 RT "Molecular cloning and expression of a third type of rabbit GDP-L-  
 RT fucose-beta-D-galactoside 2-alpha-L-fucosyltransferase.";  
 RL J. Biol. Chem. 271:16975-16981(1996).  
 CC -!- FUNCTION: CREATES A MEMBRANE-ASSOCIATED PRECURSOR OLIGOSACCHARIDE  
 CC FUC-ALPHA(1,2)GALBETA- CALLED THE H ANTIGEN WHICH IS AN  
 CC ESSENTIAL SUBSTRATE FOR THE FINAL STEP IN THE MEMBRANE-ASSOCIATED  
 CC A AND B ANTIGEN SYNTHESIS PATHWAY. H AND SE ENZYMS FUCOSYLATE THE  
 CC SAME ACCEPTOR SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.  
 CC -!- CATALYTIC ACTIVITY: GDP-L-FUCOSE + BETA-D-GALACTOSYL-R = GDP -  
 CC ALPHA-L-FUCOSYL-1,2-BETA-D-GALACTOSYL-R.  
 CC -!- PATHWAY: GLYCOSYLATION.  
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND  
 CC FORM IN TRANS CISTERNAE OF GOLGI.  
 CC -!- TISSUE SPECIFICITY: ADULT GASTROINTESTINAL TRACT AND COLON.  
 CC -!- MISCELLANEOUS: THERE ARE THREE GENES (FUT1, FUT2 AND FUT3) WHICH  
 CC ENCODE GALACTOSIDE 2-L-FUCOSYLTRANSFERASE IN RABBIT. THEY ARE  
 CC EXPRESSED IN A TISSUE-SPECIFIC MANNER.  
 CC -!- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN  
 CC GLYCOSYLTRANSFERASES.  
 CC EMBL; X91269; CAA62669.1; -.  
 DR InterPro: IPR002516; GT.11.  
 DR Pfam: PF01531; Glyco.transf\_11; 1.  
 KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;  
 KW Signal-anchor; Golgi stack.  
 FT DOMAIN 1 5  
 FT TRANSMEM 6 28  
 FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT LUMENAL, CATALYTIC (POTENTIAL).  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT SEQUENCE 347 AA; 39469 MW; 746F7007309862A5 CRC64;  
 Query Match 73.4%; Score 1391.5; DB 6; Length 347;  
 Best Local Similarity 79.8%; Pred. No. 3.2e-118;  
 Matches 260; Conservative 25; Mismatches 34; Indels 7; Gaps 1;  
 QY 1 LQQRIVKQLPSLSEKELPMTQMSSNGNTSEPMKRRDSEQHNGELRGMTINSIGRLGNQM 60  
 DB 28 LQQLRVRIQPTWEELPAL-----TPAVTRPTSQRAPSRPLGSMWTINAMRLGNQM 80  
 QY 61 GEYATLFLALARMNGRLAFIPASMHNALAFIPRISLPVLSHSDTAKKIPWQNYHLNDWMEER 120  
 DB 81 GEYATLYALAKENGRPAKIPQAQMHSTLAFIPRISLPVLSHSDTAKKIPWQNYHLNDWMEER 140  
 QY 121 YRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 180  
 DB 141 YRHIPAPVRLTGYPCSWTFYHHLRHEILKEFTLHDHVREEAQAFRLGLRVNGSRPSTFV 200  
 QY 181 GYVRRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMARCENINAS 240  
 DB 201 GYVRRGDYVHVMPQVWKGVADRGYLEKALDMFRARYSSPVFVTSNGMARCENIDAS 260  
 QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTMTGTGFWAAAYLAGGDTIYLANITLDPDSPF 300  
 DB 261 RGDVVFAGNGLESPAKDFALLTQVNTHTMTGTGFWAAAYLAGGDTIYLANITLDPDSPF 320  
 QY 301 LKVKPEAAFLPEWVGIPADISPLLK 326  
 DB 321 LKIFKPEAAFLPEWVGINADLSPLLK 346

RESULT 15  
 O77487  
 AC O77487 PRELIMINARY; PST; 343 AA.  
 DT 01-NOV-1998 (TREMELREL. 08, Created)  
 DT 01-NOV-1998 (TREMELREL. 08, Last sequence update)  
 DT 01-MAR-2002 (TREMELREL. 20, Last annotation update)  
 DE Alpha(1,2) fucosyltransferase.  
 GN FUT2.  
 OS Pongo pygmaeus (Orangutan).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pongo.  
 OX NCBI\_TaxID=9600;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;  
 RT "The old origin of a null allele sc428 of the human ABO-secretor type  
 RT alpha(1,2) fucosyltransferase gene (FUT2).";  
 RI Submitted (JUN-1998) to the EMBL/GenBank/DBSJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95181460; PubMed=7876235;  
 RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;  
 RT "Sequence and expression of a candidate for the human secretor blood  
 RT group alpha(1,2) fucosyltransferase gene (FUT2).";  
 RL J. Biol. Chem. 270:4640-4649(1995).  
 DR EMBL; AB015636; BAA31129.1; -.  
 DR InterPro: IPR002516; GT.11.  
 DR Pfam: PF01531; Glyco.transf\_11; 1.  
 KW Glycosyltransferase; Transferase.  
 SQ SEQUENCE 343 AA; 38973 MW; 2BD4D2A9704E4A0C CRC64;  
 Query Match 73.3%; Score 1389.5; DB 6; Length 343;  
 Best Local Similarity 79.1%; Pred. No. 4.8e-118;  
 Matches 258; Conservative 27; Mismatches 30; Indels 11; Gaps 2;  
 QY 1 LQQRIVKQLPSLSEKELPMTQMSSNGNTSEPMKRRDSEQHNGELRGMTINSIGRLGNQM 60  
 DB 28 LQQLRAKIQAM--WELP-----EQIPVLASTSKALGPSOLRGITWTINAILGNQM 76  
 QY 61 GEYATLFLALARMNGRLAFIPASMHNALAFIPRISLPVLSHSDTAKKIPWQNYHLNDWMEER 120  
 DB 77 GEYATLYALAKENGRPAKIPQAQMHSTLAFIPRISLPVLSHSDTAKKIPWQNYHLNDWMEER 136  
 QY 121 YRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 180  
 DB 137 YRHIPGEYVRLTGYPCSWTFYHHLRHEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 196  
 QY 181 GYVRRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMARCENINAS 240  
 DB 197 GYVRRGDYVHVMPKVKGVADRGYLEKALDMFRARYSSPVFVTSNGMARCENIDAS 256  
 QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTMTGTGFWAAAYLAGGDTIYLANITLDPDSPF 300  
 DB 257 HSDVVFAGNGIEGSPAKDFALLTQCNTHTMTGTGFWAAAYLAGGDTIYLANITLDPDSPF 316  
 QY 301 LKVKPEAAFLPEWVGIPADLSPLLK 326  
 DB 317 LKIFKPEAAFLPEWVGIAADLSPLLK 342  
 Search completed: May 27, 2003, 15:10:46  
 Job time : 38.6369 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 10:20:58 ; Search time 3098.72 Seconds  
(without alignments)  
10791.256 Million cell updates/sec  
Title: US-10-040-863-7  
Perfect score: 1149  
Sequence: 1 atgcccagcgcaggttcc.....gaagagccttctgatggaa 1149

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vl.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vl.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_higo\_hum.\*

40: em\_higo\_mus.\*

41: em\_htg\_other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1149	100.0	1149	10	AF264005	AF264005 Rattus no
2	1133.8	98.7	1847	10	AF131238	AF131238 Rattus no
3	1133.8	98.7	2984	10	AB006138	AB006138 Rattus no
4	1121.2	97.6	150840	2	AC121211	AC121211 Rattus no
5	1068	93.0	1068	10	AF042743	AF042743 Rattus no
6	931.8	81.1	6762	10	AF214656	AF214656 Mus muscu
7	931.8	81.1	200146	2	AC073774	AC073774 Mus muscu
8	865.8	75.4	1044	10	AF064792	AF064792 Mus muscu
9	802	69.8	963	10	AB039114	AB039114 Mus muscu
10	802	69.8	963	10	AB039115	AB039115 Mus muscu
11	802	69.8	963	10	AB039116	AB039116 Mus muscu
12	802	69.8	963	10	AB039117	AB039117 Mus muscu
13	802	69.8	963	10	AB039118	AB039118 Mus muscu
14	802	69.8	963	10	AB039119	AB039119 Mus muscu
15	802	69.8	963	10	AB039120	AB039120 Mus muscu
16	802	69.8	963	10	AB039121	AB039121 Mus spici
17	800.4	69.7	963	10	AB039121	AB039121 Mus muscu
18	798.8	69.5	963	10	AB039122	AB039122 Mus muscu
19	669.6	58.3	1555	10	AF131239	AF131239 Rattus no
20	659.6	57.4	1107	10	AB039213	AB039213 Mus spici
21	658	57.0	1107	10	AB039209	AB039209 Mus muscu
22	654.8	57.0	1107	10	AB039212	AB039212 Mus muscu
23	654.8	57.0	1107	10	AF113532	AF113532 Mus muscu
24	654.8	57.0	1589	10	MMSE21	Y09882 M.musculus
25	653.2	56.8	1107	10	AB039206	AB039206 Mus muscu
26	653.2	56.8	1107	10	AB039207	AB039207 Mus muscu
27	653.2	56.8	1107	10	AB039208	AB039208 Mus muscu
28	653.2	56.8	1107	10	AB039211	AB039211 Mus muscu
29	653.2	56.8	4705	10	AF214657	AF214657 Mus muscu
30	651.6	56.7	1107	10	AB039210	AB039210 Mus muscu
31	650	56.6	1107	10	AB039204	AB039204 Mus muscu
32	650	56.6	1107	10	AB039205	AB039205 Mus muscu
33	647.6	56.4	1317	4	OCAL2FT	X91269 O.cuniculus
34	633.8	55.2	2315	4	AF136895	AF136895 Sus scrofa
35	633.8	55.0	2922	4	SSU70881	U70881 Sus scrofa
36	631.8	55.0	1044	4	AF027304	AF027304 Sus scrofa
37	630.2	54.8	1043	6	AR212339	AR212339 Sequence
38	630.2	54.8	1043	6	AX029027	AX029027 Sequence
39	628.6	54.7	1043	6	AX029101	AX029101 Sequence
40	625.6	54.4	5357	10	AF214658	AF214658 Mus muscu
41	617.6	53.8	1062	9	AB015635	AB015635 Gorilla g
42	616	53.6	1062	9	AF080606	AF080606 Gorilla g
43	614.2	53.5	119631	9	AC008888	AC008888 Homo sapi
44	613.8	53.4	1784	9	BC001899	BC001899 Homo sapi
45	613.8	53.4	2115	6	AR039079	AR039079 Sequence

ALIGNMENTS

RESULT 1  
AF264005  
LOCUS  
DEFINITION  
Rattus norvegicus alpha 1-2 fucosyltransferase mRNA, complete cds.  
ACCESSION  
AF264005  
VERSION  
AF264005.1 GI:8099676  
KEYWORDS  
SOURCE  
Rattus norvegicus.  
ORGANISM  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE  
1 (bases 1 to 1149)  
AUTHORS  
Sherwood,A.L., Stroud,M.R., Levery,S.B. and Holmes,E.H.  
TITLE  
An amino acid region at the N-terminus of rat hepatoma alpha1-->2

Pred. No. is the number of results predicted by chance to have a



		/organism="Rattus norvegicus"		
		/strain="SDIX"		
		/db_xref="taxon:10116"		
		/chromosome="1"		
		/map="1q22-q31"		
		<211..>1275		
gene		/gene="FTB"		
mRNA		<211..>1275		
		/gene="FTB"		
		/product="alpha 1,2-fucosyltransferase B"		
		211..1275		
		/gene="FTB"		
		/note="glycosyltransferase; alpha 2-fucosyltransferase B"		
		/codon_start=1		
		/product="alpha 1,2-fucosyltransferase B"		
		/protein_id="AAP24469.1"		
		/db_xref="GI:4580673"		
		/translation="MASAQQPSFPLAHLLELVFVFTSTIHLQORIVKLQPISEKELP MTQMSNGTESPEMRKDSNGELGRMFTINSIGRUGNGMGEYATILFALRMNGR LAFIPASMINALAFIRISLPVLHSDTKATIPWYNHLDNMDEERYHTLPGEFVRTG YPCSWTFYHLHRELLKEPTLHDHVRREAQALRLGRVNGSQPSFTFVGVEVRGGYVH VPMWVKVQDRGLYLEKALDEMRARYSSPVFVTTSGMAWCRENINASRGDGVFAGN GIEGSPAKDFALLTQCNHNTIMTIGTFIWAAYLAGDITVILANYTLPDSFFLLKVPKPE AAFLPEWGVIPADLSPLKH"		
BASE COUNT	452 a	503 c	457 g	435 t
ORIGIN				
	Query Match	98.7%;	Score 1133.8;	DB 10; Length 1847;
	Best Local Similarity	99.7%;	Pred. No. 9.3e-293;	
	Matches 1146;	Conservative	0;	Mismatches 2; Indels 1; Gaps 1;
QY	1	ATGGCCAGCGCCGAGTTCTCTTCCTTTCTCCTTTCTCGCCACCTTCCTCATCTTTGCTTC	60	
Db	211	ATGGCCAGCGCCGAGTTCTCTTCCTTTCTCCTTTCTCGCCACCTTCCTCATCTTTGCTTC	270	
QY	61	GTGACTTCCACCATCATCACCTCCAGCAGCGNATAGTGAAGCTTCAACCCCTGTGAG	120	
Db	271	GTGACTTCCACCATCATCACCTCCAGCAGCGAATAGTGAAGCTTCAACCCCTGTGAG	330	
QY	121	AAGGAATTACCGATGACGACTCAAAATGTCTCGGGAACACAGAAAGCCACAGATGCG	180	
Db	331	AAGGAATTACCGATGACGACTCAAAATGTCTCGGGAACACAGAAAGCCACAGATGCG	390	
QY	181	CGGACAGCGACGACGATGGGAATGGAGAGCTCGGGGCGATGTTACGATCAATTCATT	240	
Db	391	CGGACAGCGAGCAGCATGGGAATGGAGAGCTCGGGGCGATGTTACGATCAATTCATT	450	
QY	241	GGCCGCGTGGGNACGATGGCGGAATAGCCACACTCTTTGCACTGGCAGGATGAAC	300	
Db	451	GGCCGCGTGGGAACACAGATGGCGGAATAGCCACACTCTTTGCACTGGCAGGATGAAC	510	
QY	301	GGACGCGTTGGCTTCATCCCGCATCCATGCACACGCTCTAGCGCCCATCTTCAGGATC	360	
Db	511	GGACGCGTTGGCTTCATCCCGCATCCATGCACACGCTCTAGCGCCCATCTTCAGGATC	570	
QY	361	AGCCTCCCGGTGTTACACAGCGACACGGCCAAAAGATCCCATGGCAGAATTACCATCTC	420	
Db	571	AGCCTCCCGGTGTTACACAGCGACACGGCCAAAAGATCCCATGGCAGAATTACCATCTC	630	
QY	421	AACGACTGGATGGAGAGCGTTACCGCCACATCCGGGACACTTTGTGGCTTCACGGGA	480	
Db	631	AACGACTGGATGGAGAGCGTTACCGCCACATCCGGGACACTTTGTGGCTTCACGGGA	690	
QY	481	TACCGTGTCTTGGACCTTCTACCAACACCTCGGCCACAGATCTCGAAGGAGTTCAACC	540	
Db	691	TACCGTGTCTTGGACCTTCTACCAACACCTCGGCCACAGATCTCGAAGGAGTTCAACC	750	
QY	541	CTGCATACACAGTGGCGGAGGAGGCCCGCTTCTCGTGGTCTGGCGGGTGAATGGG	600	
Db	751	CTGCATACACAGTGGCGGAGGAGGCCCGCTTCTCGTGGTCTGGCGGGTGAATGGG	810	
QY	601	AGCCACCGAGTACTTTGTGGGTGTCCATGTGCGCCGAGGGGACTATGTGCATGTCAAT	660	

Db	111	ASCCAGCCGAGTACTTTTGTGGGTGTCCATCTGCGCGAGGGGACTATGTCCATGTGATG 870
QY	661	CCTAATGTGTGGAAGGCGTGGTGGCTGACCGCGGGTTACCTGGAAAAGGCCCTGGATATG 720
Db	871	CTTAATGTGTGGAAGGCGTGGTGGCTGACCGCGGGTTACCTGGAAAAGGCCCTGGATATG 930
QY	721	TTCCGGGSCAGCTTAATCATCTCCAGTCTTCGTGGTTACAAGCAACGGTATGGCCTGGTGC 780
Db	931	TTCCGGGSCAGCTTAATCATCTCCAGTCTTCGTGGTTACAAGCAACGGTATGGCCTGGTGC 990
QY	781	CGGAGCAACATTAATGCTTCCCGAGGAGAGCTGGTGTTCGCGGGCAATGTAATGAGGG 840
Db	991	CGGAGCAACATTAATGCTTCCCGAGGAGAGCTGGTGTTCGCGGGCAATGTAATGAGGG 1050
QY	841	TCGCCAGCAAGGACTTCGCGGTGCTCACCCAGTGCAACACACACACCATCATGACTATTGGG 900
Db	1051	TCGCCAGCAAGGACTTCGCGGTGCTCACCCAGTGCAACACACACCATCATGACTATTGGG 1110
QY	901	ACCTTTGGGATTTGGCTGCTACCTACCTGGCAGGTGGTGATACCATCTACTTAGCCAACTAC 960
Db	1111	ACCTTTGGGATTTGGCTGCTACCTACCTGGCAGGTGGTGATACCATCTACTTAGCCAACTAC 1170
QY	961	ACCTTTCCGATCTCCGTTCTCCCAAGTCTTTAAGCAGAGGAGGCTTCTTACCGGAA 1020
Db	1171	ACCTTTCCGATCTCCGTTCTCCCAAGTCTTTAAGCAGAGGAGGCTTCTTACCGGAA 1230
QY	1021	TGGTGGGATCCCTGCGGATCTGCTCCCACTCTTAAGGCAATTAACACAGGCTGTCCCT 1080
Db	1231	TGGTGGGATCCCTGCGGATCTGCTCCCACTCTTAAGGCAATTAACACAGGCTGTCCCT 1289
QY	1081	CGGTCCCACTTCCACTCAAGGCAAGGAGTCACTTGTACGTCGAGGAGGAGGCTTTC 1140
Db	1290	CGGTCCCACTTCCACTCAAGGCAAGGAGTCACTTGTACGTCGAGGAGGAGGCTTTC 1349
QY	1141	TGATGGGAA 1149
Db	1350	TGATGGGAA 1358
RESULT 3		
AB006138		
LOCUS	AB006138	2984 bp mRNA linear ROD 05-FEB-1999
DEFINITION	Rattus norvegicus FTB mRNA for alpha 1,2-fucosyltransferase,	
	complete cds.	
ACCESSION	AB006138	
VERSION	AB006138.1	GI:2317265
KEYWORDS	FTB; alpha 1,2-fucosyltransferase.	
SOURCE	Rattus norvegicus	cell_line:RCN-9 cDNA to mRNA.
ORGANISM	Rattus norvegicus	
REFERENCE	1 (sites)	
AUTHORS	Soejima,M., Wang,B., Koda,Y. and Kimura,H.	
TITLE	Two distinct rat GSP-L-fucose:b-D-galactoside	
JOURNAL	2-a-L-fucosyltransferase genes	
AUTHORS	Unpublished	
TITLE	2 (bases 1 to 2984)	
JOURNAL	Koda,Y.	
REFERENCE	Submitted (04-AUG-1997)	
AUTHORS	Yoshiro Koda, Kurume University, School of	
TITLE	Medicine, Department of Forensic Medicine; Asahimachi 67, Kurume,	
JOURNAL	Fukuoka 830, Japan (E-mail:ykoda@med.kurume-u.ac.jp,	
REFERENCE	Tel:0942-31-7554, Fax:0942-31-7700)	
FEATURES	Location/Qualifiers	
source	1..2984	
	/organism="Rattus norvegicus"	
	/db_xref="taxon:10116"	
	/chromosome="1"	
	/map="1q22-q31"	
	/cell_line="RCN-9"	
	/tissue_type="colon cancer"	
gene	1..2984	

CDS	/gene="FTB" 214..1278 /gene="FTB" /codon_start=1 /product="alpha 1,2-fucosyltransferase" protein_id="BAA21742.1" /db_xref="GI:2317266" /translation="MASAQVFPSPPLAHFLIFVFVSTIIHLQORIVKIQPSEKELPMTTMSNTSPEMRSDSHGNGSELGRMFTINSTIRGLNOMGEYATLFAALMMNRGLAFIPASNNALADIFAISILPFDSTAKKIPWNYHLDNMEEYRHLPGFVRPTGYPCSWTFVHLRPLIKFEFLHDHVREEQAFLRG.FRVNGSQSFSTFVGVRGGDYVHVMPNVKGVADRGLYLEKALIDMPARYSSVPVVTSNGMAWCRENTNASRGDVPFRAGNGIESPKADFALLTCQNHITLMVIGTGFOINRAAYLAGDTIYLANITYLPDSPFLKVPKEAAFIPEWGVIADPLSPLLKH"
BASE COUNT	729 a 808 c 761 g 686 t
ORIGIN	
	Query Match 98.7%; Score 1133.8; DB 10; Length 2984; Best Local Similarity 99.7%; Pred. No. 9.5e-293; Matches 1146; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY	1 ATGGCCAGCGCCAGGTTCCTTTCTCCTTCTCTCTGCGCCACATTCCTCATCTTTTGCTTC 60
Db	214 ATGGCCAGCGCCAGGTTCCTTTCTCCTTCTCTCTGCGCCACATTCCTCATCTTTTGCTTC 273
QY	61 GTGACTTCCACCATCATCACTCCAGCAGCGGAATAGTGAAGCTCCAACCCCTGCAGAG 120
Db	274 GTGACTTCCACCATCATCACTCCAGCAGCGGAATAGTGAAGCTCCAACCCCTGCAGAG 333
QY	121 AAGGAATTACCGATGACGACTCAAATGCTCTCGGGAAACACAGAAAACCCAGAGATCGGA 180
Db	334 AAGGAATTACCGATGACGACTCAAATGCTCTCGGGAAACACAGAAAACCCAGAGATCGGA 393
QY	181 CGSGACAGCAGCACATGGGAATGGAGAGCTCGGGSCATGTTCAGCATCAATTCCTATT 240
Db	394 CGGGACAGCAGCACATGGGAATGGAGAGCTCGGGSCATGTTCAGCATCAATTCCTATT 453
QY	241 GCCCGGCTGGGAAACAGATGGGGAATACGCCACACTTTTGACATGGCCAGAGATGAAC 300
Db	454 GSCCGGCTGGGAAACAGATGGGGAATACGCCACACTTTTGACATGGCCAGAGATGAAC 513
QY	301 GAACGGCTTGGTTTCATCCCGCATCCATGCACAACGCCTTAGCGCCCATCTTCAGGATC 360
Db	514 GAACGGCTTGGTTTCATCCCGCATCCATGCACAACGCCTTAGCGCCCATCTTCAGGATC 573
QY	361 AGCCTCCCGGTGTTACACAGCGAACAGGCCAAAAAGATCCCATGGCAGAATTACCATCTC 420
Db	574 ASCCTCCCGGTGTTACACAGCGAACAGGCCAAAAAGATCCCATGGCAGAATTACCATCTC 633
QY	421 AACGACTGGATGGAGAGCGTTACGGCCAAATTCGGGCACACTTTGTGCGCTTCACGGGA 480
Db	634 AACGACTGGATGGAGAGCGTTACGGCCAAATTCGGGCACACTTTGTGCGCTTCACGGGA 693
QY	481 TACCCTGTCTCTGCACTTCTACACCACTCGCCGACAGATCCCTGAAGGAGTTCACC 540
Db	694 TACCCTGTCTCTGCACTTCTACCACTCGCCGACAGATCCCTGAAGGAGTTCACC 753
QY	541 CTGCTATGACCACTGCGGGAGAGCCCAAGCCCTTCTCGGTGTCTGCGGGTGAATGGG 600
Db	754 CTGCTATGACCACTGCGGGAGAGCCCAAGCCCTTCTCGGTGTCTGCGGGTGAATGGG 813
QY	601 ACCCAGCCGAGTACTTTGTGGTGTCTCCATGTGCGCGGAGGACIATGTGCATGTCATG 660
Db	814 ACCCAGCCGAGTACTTTGTGGTGTCTCAATGTGCGCGGAGGACIATGTGCATGTCATG 873
QY	661 CTAATGTGTGAAGGCGGTGTGCTGACCGGGTGTACCTTGGAAAAGCCCTGGATATG 720
Db	874 CTAATGTGTGAAGGCGGTGTGCTGACCGGGTGTACCTTGGAAAAGCCCTGGATATG 933
QY	721 TTCCGGGACCGCTATTCATCTCCAGTCTTCGTGTTACACACGATGGCTATGGCTGGTGC 780
Db	934 TTCCGGGACCGCTATTCATCTCCAGTCTTCGTGTTACACACGATGGCTATGGCTGGTGC 993
QY	781 CGGAGAACATAATTAATGCTTCCCGAGGAGACGTGGTGTTCGCGGGCAATGCTATTGAGGG 840
Db	994 CGGAGAACATAATTAATGCTTCCCGAGGAGACGTGGTGTTCGCGGGCAATGCTATTGAGGG 1053
QY	841 TCGCCAGCAAGGACTTCGGCTCTCACCCAGGCAACACACACCACTCATGACTATTGGG 900
Db	1054 TCGCCAGCAAGGACTTCGGCTCTCACCCAGGCAACACACACCACTCATGACTATTGGG 1113
QY	901 ACCTTTGGGATTTGGGCTGCTACCTGGCAGGTGGTGATACCATCTACTTTAGCCAACTAC 960
Db	1114 ACCTTTGGGATTTGGGCTGCTACCTGGCAGGTGGTGATACCATCTACTTTAGCCAACTAC 1173
QY	961 ACCCTTCGGATTCCTCGTTCTCAAAAGTCTTTAAGCCAGCAGGACGCTTCCTACCCGAA 1020
Db	1174 ACCCTTCGGATTCCTCGTTCTCAAAAGTCTTTAAGCCAGCAGGACGCTTCCTACCCGAA 1233
QY	1021 TGGGTGGGCATCCCTGCGGATCTGCCCACTCTCTTAAGGCATTAAACACAGGCTCTCT 1080
Db	1234 TGGGTGGGCATCCCTGCGGATCTGCCCACTCTCTCTTAAGGCATTAAACACAGGCTCTCT 1292
QY	1081 CGGTCCCACCTTCCACCTCAAGGCAAAAGGAGTCACTTTGTTACGTGCGAGGAAGGCTTC 1140
Db	1293 CGGTCCCACCTTCCACCTCAAGGCAAAAGGAGTCACTTTGTTACGTGCGAGGAAGGCTTC 1352
QY	1141 TGATCGGAA 1149
Db	1353 TGATCGGAA 1361
RESULT 4	
AC121211	Rattus norvegicus clone CH230-309A14, *** SEQUENCING IN PROGRESS
LOCUS	*** 63 unordered pieces.
DEFINITION	AC121211.2 GI:21909463
ACCESSION	HTG: HTGS_PHASE1.
VERSION	Rattus norvegicus.
KEYWORDS	Rattus norvegicus.
SOURCE	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;
ORGANISM	Rattus.
REFERENCE	1 (bases 1 to 150840)
AUTHORS	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Avele,M., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Briefa,M., Brown,E., Brown,M., Bryant,N.P., Bubay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.G., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,I., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haviak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,B., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratochvic,D., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichteberg,O., Lieu,C., Liu,J., Liu,W., Louisedge,H., Lozaro,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mathewine,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,S., Nguyen,N., Nguyen,A., Nguyen.N., Nguyen.N., Nickerson,E., Nwokendo,S., Oghu,M., Okwuono,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L., Quiles,M., Ren,Y.,



Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G.,  
 Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,  
 Sodergren, E., Sonaika, J., Sparks, A., Stanley, H., Stone, I.,  
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
 Tansey, J., Taylor, C., Taylor, I., Teifrod, B., Thomas, N., Thomas, S.,  
 Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
 Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,  
 Wu, C., Wu, Y., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G. and Gibbs, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 150840)  
 Worley, K.C.  
 Direct Submission  
 Submitted (16-MAY-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 150840)  
 Worley, K.C.  
 Direct Submission  
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 19, 2002 this sequence version replaced gi:20806241.

## COMMENT

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GYKK  
 Center clone name: CH230-309A14  
 ----- Summary Statistics  
 Sequencing vector: Plasmid  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 106837 bases at least Q40  
 Consensus quality: 113151 bases at least Q30  
 Consensus quality: 116417 bases at least Q20  
 -----

\* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 63 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 1027: contig of 1027 bp in length  
 1028 1127: gap of unknown length  
 1128 2276: contig of 1149 bp in length  
 2277 2376: gap of unknown length  
 2377 3792: contig of 1416 bp in length  
 3793 3892: gap of unknown length  
 3893 4898: contig of 1006 bp in length  
 4899 4999: gap of unknown length  
 5000 6224: contig of 1226 bp in length  
 6225 6325: gap of unknown length  
 6326 7677: contig of 1353 bp in length  
 7678 7777: gap of unknown length  
 7778 8779: contig of 1002 bp in length  
 8780 8879: gap of unknown length  
 8880 10342: contig of 1463 bp in length  
 10343 10442: gap of unknown length  
 10443 11594: contig of 1152 bp in length  
 11595 11694: gap of unknown length  
 11695 13897: contig of 2203 bp in length  
 13898 13998: gap of unknown length  
 13999 15482: contig of 1485 bp in length  
 15483 15582: gap of unknown length

15583 17010: contig of 1428 bp in length  
 17011 17110: gap of unknown length  
 17111 18427: contig of 1317 bp in length  
 18428 18527: gap of unknown length  
 18529 19811: contig of 1284 bp in length  
 19812 21647: contig of 1736 bp in length  
 21648 22753: contig of 1006 bp in length  
 22754 22854: gap of unknown length  
 22855 24607: contig of 1753 bp in length  
 24608 26167: contig of 1461 bp in length  
 26168 26267: gap of unknown length  
 26268 27443: contig of 1176 bp in length  
 27444 27543: gap of unknown length  
 27544 28739: contig of 1196 bp in length  
 28740 28839: gap of unknown length  
 28840 30991: contig of 2152 bp in length  
 30992 32113: contig of 1022 bp in length  
 32114 33372: contig of 1159 bp in length  
 33373 33472: gap of unknown length  
 33473 35105: contig of 1633 bp in length  
 35106 35205: gap of unknown length  
 35206 36378: contig of 1173 bp in length  
 36379 36478: gap of unknown length  
 36479 38090: contig of 1612 bp in length  
 38091 38190: gap of unknown length  
 38191 39491: contig of 1301 bp in length  
 39492 41132: contig of 1541 bp in length  
 41133 41232: gap of unknown length  
 41233 42973: contig of 1741 bp in length  
 42974 43073: gap of unknown length  
 43074 44652: contig of 1579 bp in length  
 44653 44752: gap of unknown length  
 44753 45870: contig of 1118 bp in length  
 45871 45970: gap of unknown length  
 45971 48400: contig of 2430 bp in length  
 48401 48500: gap of unknown length  
 48501 50342: contig of 1842 bp in length  
 50343 50442: gap of unknown length  
 50443 52339: contig of 1897 bp in length  
 52340 52439: gap of unknown length  
 52440 53860: contig of 1421 bp in length  
 53861 53960: gap of unknown length  
 53961 56447: contig of 2487 bp in length  
 56448 56547: gap of unknown length  
 56548 57889: contig of 1342 bp in length  
 57890 57989: gap of unknown length  
 57990 59213: contig of 1224 bp in length  
 59214 59313: gap of unknown length  
 59314 61585: contig of 2272 bp in length  
 61586 61685: gap of unknown length  
 61686 63789: contig of 2104 bp in length  
 63790 63889: gap of unknown length  
 63890 65508: contig of 1619 bp in length  
 65509 65608: gap of unknown length  
 65609 67142: contig of 1534 bp in length  
 67143 67242: gap of unknown length  
 67243 69536: contig of 2294 bp in length  
 69537 69636: gap of unknown length  
 69637 72776: contig of 3140 bp in length  
 72777 72876: gap of unknown length  
 72877 74143: contig of 1267 bp in length  
 74144 74243: gap of unknown length  
 74244 77710: contig of 3467 bp in length  
 77711 77810: gap of unknown length  
 77811 80194: contig of 2384 bp in length  
 80195 80294: gap of unknown length  
 80295 82739: contig of 2445 bp in length



Best Local Similarity 100.0%; Pred. No. 3.9e-275;  
Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	82	CTCCAGCAGCAAGTGTGAAGCTCCAAACCCCTGTGCAGAGAGGAATTACCGATGACGACT	141
Db	1	CTCCAGCAGCAAGTGTGAAGCTCCAAACCCCTGTGCAGAGAGGAATTACCGATGACGACT	60
QY	142	CAATGTCTCTCGGGAACACAGAAACCCAGAGATCGACGCGAGCAGCAGCATGGG	201
Db	61	CAATGTCTCTCGGGAACACAGAAACCCAGAGATCGACGCGAGCAGCAGCATGGG	120
QY	202	RATGAGAGCTGCGGGGATGTTACAGATCAATTCATTGCGCGGCTGGGAACAGATG	261
Db	121	RATGAGAGCTGCGGGGATGTTACAGATCAATTCATTGCGCGGCTGGGAACAGATG	180
QY	262	GCGAATACGCCACACTCTTGCACTGGCCAGGATGAACGGACGCTTGCCTTCATCCCC	321
Db	181	GCGAATACGCCACACTCTTGCACTGGCCAGGATGAACGGACGCTTGCCTTCATCCCC	240
QY	322	GCATCCATGCAACAGCTCTAGCGCCCATCTTCAGGATCAGCCTCCCGTGTACACAGC	381
Db	241	GCATCCATGCAACAGCTCTAGCGCCCATCTTCAGGATCAGCCTCCCGTGTACACAGC	300
QY	382	GACAGGCGCAAAAGATCCCATGGCAGATTAACATCTCAACGATGGATGAGGAGCGT	441
Db	301	GACAGGCGCAAAAGATCCCATGGCAGATTAACATCTCAACGATGGATGAGGAGCGT	360
QY	442	TACGCCACATTCGGGGACACTTGTGGCTGTACGGGATACCGGTCTCTGGACCTTC	501
Db	361	TACGCCACATTCGGGGACACTTGTGGCTGTACGGGATACCGGTCTCTGGACCTTC	420
QY	502	TACCACACCTCGCGCCAGAGATCCTGAAGGAGTTTCAACCTGCTCATSACACAGTGGGGAG	561
Db	421	TACCACACCTCGCGCCAGAGATCCTGAAGGAGTTTCAACCTGCTCATSACACAGTGGGGAG	480
QY	562	GAGGCCAGGCTTCTCGTGGTCTGGGGTGAATGGAGCCAGCCAGTACTTTGTG	621
Db	481	GAGGCCAGGCTTCTCGTGGTCTGGGGTGAATGGAGCCAGCCAGTACTTTGTG	540
QY	622	GTTGTCCATGTGCGCGAGGAGCACTATGTGATGTCATGCTAATGTGGAGGCGTG	681
Db	541	GTTGTCCATGTGCGCGAGGAGCACTATGTGATGTCATGCTAATGTGGAGGCGTG	600
QY	682	GTGGTGAACCGGGTTACCTGGAAGCCCTGGATATGTTCCGGGCAACGCTTATTCATCT	741
Db	601	GTGGTGAACCGGGTTACCTGGAAGCCCTGGATATGTTCCGGGCAACGCTTATTCATCT	660
QY	742	CCAGTCTTCGTGGTTACAGCAACGGTATGGCTGGTCCCGGGAGAACATTAATGCTTCC	801
Db	661	CCAGTCTTCGTGGTTACAGCAACGGTATGGCTGGTCCCGGGAGAACATTAATGCTTCC	720
QY	802	CGAGGAGACGTGGTGTTCGCGGGCAATGTTAGGGGTGCGCCAGCAAGCACTTCGCG	861
Db	721	CGAGGAGACGTGGTGTTCGCGGGCAATGTTAGGGGTGCGCCAGCAAGCACTTCGCG	780
QY	862	CTGCTCACCCAGTGAACCAACCACTATGACTATGGGACCTTTGGGATTTGGGCTGCC	921
Db	781	CTGCTCACCCAGTGAACCAACCACTATGACTATGGGACCTTTGGGATTTGGGCTGCC	840
QY	922	TACCTGGGAGTGTGATACCATCTACTTAGCCACTACACCTTCGGGATTCCTGCTC	981
Db	841	TACCTGGGAGTGTGATACCATCTACTTAGCCACTACACCTTCGGGATTCCTGCTC	900
QY	982	CTCAAAGTCTTTAAGCCAGAGCAGCCTTCCTACCCGAATGGTGGGATCCCTCCCGAT	1041
Db	901	CTCAAAGTCTTTAAGCCAGAGCAGCCTTCCTACCCGAATGGTGGGATCCCTCCCGAT	960
QY	1042	CTGTCCCACTCTTTAAGGATTAACACAGCCTTCTCTCGTCCCACTTCCACCTCAAG	1101
Db	961	CTGTCCCACTCTTTAAGGATTAACACAGCCTTCTCTCGTCCCACTTCCACCTCAAG	1020
QY	1102	GCAAGAGAGTCACTTGTAGTCCAGAGAGCCTTCTGATGGGA	1149

Db	1021	GCAAGAGAGTCACTTGTAGTCCAGAGAGCCTTCTGATGGGA	1068
RESULT 6			
AF214656		6752 bp	DNA linear ROD 26-JUN-2001
LOCUS			
DEFINITION			Mus musculus alpha(1,2)fucosyltransferase FUT2 (Fut2) gene, complete cds.
ACCESSION	AF214656		
VERSION	AF214656.1	GI:7288504	
KEYWORDS			
SOURCE			Mus musculus.
ORGANISM			Mus musculus
REFERENCE			1 (bases 1 to 6762)
AUTHORS			Domino,S.E., Zhang,L. and Lowe,J.B.
TITLE			Molecular cloning, genomic mapping, and expression of two secretor blood group alpha (1,2)fucosyltransferase genes differentially regulated in mouse uterine epithelium and gastrointestinal tract
JOURNAL			J. Biol. Chem. 276 (26), 23748-23756 (2001)
MEDLINE			21316545
PUBMED			11323419
REFERENCE			2 (bases 1 to 6762)
AUTHORS			Domino,S.E. and Lowe,J.B.
TITLE			Mus musculus alpha(1,2)fucosyltransferase FUT2 (Sec2) genomic sequence
JOURNAL			Unpublished
AUTHORS			3 (bases 1 to 6762)
TITLE			Domino,S.E. and Lowe,J.B.
JOURNAL			Direct Submission
FEATURES			Submitted (13-DEC-1999) HHMI, University of Michigan, 1150 W. Med. Ctr. Dr., Ann Arbor, MI 48109-0650, USA
source			Location/Qualifiers
			1..6762
			/organism="Mus musculus"
			/strain="129/Ola"
			/db_xref="taxon:10090"
			/chromosome="7"
			/map="23.2: between k1k1 and Ldh3"
			/cell_line="E14"
			/cell_type="ES"
			<2210..>3253
			/gene="Fut2"
			/note="Sec2"
			<2210..>3253
			/gene="Fut2"
			/product="alpha(1,2)fucosyltransferase FUT2"
			2210..3253
			/gene="Fut2"
			/EC_number="2.4.1.69"
			/note="Fucosyltransferase: similar to the Mus musculus strain ICR Sec2 sequence deposited at GenBank Accession Number AF064792; similar to the Homo sapiens secretor blood group alpha(1,2)fucosyltransferase FUT2 sequences deposited at GenBank Accession Numbers D87942 and U17894"
			/codon_start=1
			/product="alpha(1,2)fucosyltransferase FUT2"
			/protein_id="AAF45146.1"
			/db_xref="GI:7288505"
			/translation="MASAQVPFSPFLAHFLIFVFVSTIIHQORIVKLOTLSKELQAVQMSPPAARTDQQSAKLAGIFITNSIGRLNQMGYATLFLARVNGRLAIPESMHNALAFRISLPVLHSDTARRIPWONHLNDWMEERYRHTPGGVVRETPGCSWTFYHLRPLELHKEFTLDHVREREAQATLRLGRVNGSOPSTFVGHVHVRGDIYVMPKVKGVADRGVLEKALDRFRARYSSPYVTVTSNGMANKRENITSLGDVVFAGVIEGSPA KDFALLTCNHTITMTIGTFFGIWAAYLAGDGIYIANYTIPDSPFLKIFAPAAAFLEW MGIPADLSPLLK"
			polyA_signal 4943..4948
			polyA_signal 5333..5338
			BASE COUNT 1738 a 1682 c 1647 g 1691 t 4,others
			ORIGIN
			Query Match 81.1%; Score 931.8; DB 10; Length 6762;



```

* 113115 117991: contig of 4877 bp in length
* 117992 118091: gap of unknown length
* 118092 148417: contig of 28326 bp in length
* 148418 148517: gap of unknown length
* 148518 154588: contig of 8071 bp in length
* 154589 154688: gap of unknown length
* 154689 158642: contig of 3954 bp in length
* 158643 158742: gap of unknown length
* 158743 160577: contig of 1835 bp in length
* 160578 160677: gap of unknown length
* 160678 198882: contig of 36205 bp in length
* 198883 198982: gap of unknown length
* 198983 200146: contig of 3164 bp in length.
FEATURES             Location/Qualifiers
     source           1..200146
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone_lib="RP23-36F17"
                     /clone_lib="RPCI mouse BAC library 23"
BASE COUNT          47482 a 52711 c 50595 g 48057 t 1301 others
ORIGIN
Query Match          81.1%; Score 931.8; DB 2; Length 200146;
Best Local Similarity 89.7%; Pred. No. 1.5e-238;
Matches 1031; Conservative 0; Mismatches 97; Indels 21; Gaps 2;
QY      1  ATGCCACGCCCGAGGTTCCTTCTCCTTCTCTGCCCCACTTCCCTCATCTTGTCTTC 60
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      14017 ATGGGAGTGGCCAGGATACCTTCTCCTTCTCTGCCCCACTTCCCTCATCTTGTCTTC 13958
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      61  GTGACTTCCACCATCATCCACTCCAGCAGCGAATAGTGAAGCTCCCAACCCCTGTCAGAG 120
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      13957 GTGACTTCCACCATCATCCACTCCAGCAGCGAATAGTGAAGCTCCCAACCCCTGTCAGAG 13898
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      121  AAGAAATACGATGAGCATCAAAATGTCCTCGGAAACACAGAAAGCCAGAGATCGGA 180
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      13897 AAGAAATATC---AGGGGGTTTCAAATGTCTTCCACAAAGCGGCAAGAACAGACAI---- 13845
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      181  CGGCACAGCAGCAGATGGAATGGAGAGCTGCGGGGCATGTTTCAGATCAATTCGATT 240
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      13844 -----GCAGCAGATGCCAAGCTGCAGGCGCATTTACGATCAATTCATC 13799
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      241  GCGCGGTGGGAAACAGATGGGCGAATACGCCACTCTTTGCACTGGCCAGGATGAAC 300
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      13798 GCGCGCTGGGAAACAGATGGGCGAATGCTACATGTTTGCATGGCCAGATGAAC 13739
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      301  GGACGGCTGCTTCATCCCGCATCCATGACACACGCTGAGCGGCCATCTTCAGGATC 360
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      13738 GGTGGGTTCCTTTCATCCCTGAAATGATGACACACGCTCTAGCGGCCATCTTCAGGATC 13679
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      361  AGCTCCCGGTGTACACAGCGACACGCCCAAAAGATCCCATGCGAGAAATACCATCTC 420
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      13678 AGTCCCGGTGTACACAGCGACACAGCGAGAGATCCCGTGGCAGAAATACCATCTC 13619
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      421  AACGACTGGATGGAGAGCGTTTACCGGCACATTCGGGACACTTGTGCGCTTCACGGGA 480
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      13618 AACGACTGGATGGAGAGCGTTTACCGGCACATTCGGGACACTTGTGCGCTTCACGGGA 13559
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      481  TACCGGTGCTCTGACCTTCTACACCACTCTGCGCCAGAGATCTTGAAGAGATTAC 540
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      13558 TACCGGTGCTCTGACCTTCTACACCACTCTGCGCCAGAGATCTTGAAGAGATTAC 13499
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      541  CTGCATGACACGTCGCGGAGAGCGCCAGCCCTTCTGCTGCTGCTGCGGGTGAATGG 600
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      13498 CTGCAGACCATGTCGTCGAGAGCGCCAGCTTCTGCTGCTGCTGCGGGTGAATGG 13439
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      601  AGCCAGCGGAGTACTTTTGTGGGTGTCATGTCGCCGAGGGGACTATGTGCATGTCATG 660
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      13438 AGCCAGCGGAGTACTTTTGTGGGTGTCATGTCGCCGAGGGGACTATGTGCATGTCATG 13379
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      661  CCTAATGTGGAAGGCGTGGTGGCTGACCGGGGTACCTGGGAAAGGCCCTGGATG 720
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      13378 CCCAAGGTGGAAGGCGTGGTGGCTGACCGGGGTACCTGGGAAAGGCCCTGGAGAGG 13319
DB      11  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

QY      721  TTCGGGCAACGCTATTCATCTCCAGTCTTCGTGGTTACAGCAACGGTATGCCCTGGTGC 780
DB      13318 TTCGGGCAACGCTATTCATCTCCAGTCTTCGTGGTTACAGCAACGGTATGCCCTGGTGC 13259
QY      781  CGGAGAACATTAAATGCTTCCGAGGAGACGCTGGTGTTCGCGGCAATGTTATGAGGG 840
DB      13258 CGGAGAACATTAAATGCTTCCGAGGAGACGCTGGTGTTCGCGGCAATGTTATGAGGG 13199
QY      841  TCGCAGCAAGAGACTTCGCGCTGCTCCACAGTGCACCAACCAACCATCATGACTATGGG 900
DB      13198 TCGCAGCAAGAGACTTCGCGCTGCTCCACAGTGCACCAACCAACCATCATGACTATGGG 13139
QY      901  ACCTTTGGGATTTGGCTGCTTACCTGCGAGGTGTGATACCATCTACTTACCGCAACTAC 960
DB      13138 ACCTTTGGGATTTGGCTGCTTACCTGCGAGGTGTGATACCATCTACTTACCGCAACTAC 13079
QY      961  ACCCTTCGGGATTTCCCTGCTTCAAGCTTTAAGCCAGAGGAGCGCTTCTACCGCAA 1020
DB      13078 ACCCTTCGGGATTTCCCTGCTTCAAGCTTTAAGCCAGAGGAGCGCTTCTACCGCAA 13019
QY      1021  TGGGTGGGATCCCGCGGATCTGTCCCACTCTTAAAGCAATTAACACCAAGCGCTTCT 1080
DB      13018 TGGGTGGGATCCCGCGGAGCTGTCCCACTCTTAAAGCAATTAACACCAAGCGCTTCT 12959
QY      1081  CGGTCCCACTTCCCACTCAAGGCAAGAGAGTCACTTCTTACGTCGCGAGGAGCGCTTC 1140
DB      12958 TGGTCCCACTTCCCACTCAAGGCAAGAGAGTCACTTCTTACGTCGCGAGGAGCGCTTC 12899
QY      1141  TGATGGGAA 1149
DB      12898 TACTGGGAA 12890

RESULT 8
AF064792          1044 bp mRNA linear ROD 24-OCT-2000
LOCUS             AF064792
DEFINITION        Mus musculus GDP-L-fucose:beta-D-galactoside
ACCESSION         AF064792
VERSION           AF064792.1 GI:3142705
KEYWORDS          Mus musculus
SOURCE            Mus musculus
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE         1 (bases 1 to 1044)
AUTHORS           Lin,B., Hayashi,Y., Saito,M., Sakakibara,Y., Yanagisawa,M. and
                  Iwamori,M.
TITLE             GDP-fucose: beta-galactoside alpha1,2-fucosyltransferase, MFUT-II,
                  and not MFUT-I or -III, is induced in a restricted region of the
                  digestive tract of germ-free mice by host-microbe interactions and
                  cycloheximide
JOURNAL           Biochim. Biophys. Acta 1487 (2-3), 275-285 (2000)
MEDLINE          20471982
PUBMED           11018479
REFERENCE         2 (bases 1 to 1044)
AUTHORS           Lin,B., Hayashi,Y., Saito,M., Sakakibara,Y., Yanagisawa,M. and
                  Iwamori,M.
TITLE             Direct Submission
JOURNAL           Submitted (12-MAY-1998) Chemistry, Faculty of Science & Technology,
                  Kinki University, Kowakae, Higashi-Osaka 577-8502, Japan
FEATURES          Location/Qualifiers
     source        1..1044
                  /organism="Mus musculus"
                  /strain="ICR"
                  /db_xref="taxon:10090"
                  /tissue_type="gastrointestinal tract"
                  1..1044
                  /EC_number="2.4.1.69"
                  /note="sec2; Se"
                  /codon_start=1
                  /product="GDP-L-fucose:beta-D-galactoside"
CDS

```

2-alpha-L-fucosyltransferase"  
/protein\_id="AAC16887.1"  
/db\_xref="GI:3142706"  
/translation="MASAOKVPEFPLAHFLIFVFTVSTIIHLQQIVKLTSEKELO  
AVQMSNSNARTDMQSQSLQGLFTINSIGRLNOMGEYATLALANMGRALFIPESMINALAFIPRI  
MHRLAPFLKSLPLVHSDTARPIWYHLNDWMEERYRHIPQYVYFTGYPCSTF  
YHLRPLKLEFTLHDHVRBEAFLGLRVNGSQPSTFVGVHVRDGYVHVPKVKW  
GVVADRGYLEKALDRFRARYSSPVVTSNMAWCRENINTSLGDDVFAVAGNIEGSPA  
KDFALLTQCCHNTIMTIGTGIWVAYLAGGDTIVLVNTLPDPSFLKIFKFAANFLPWW  
MGIPADLSPLLKH"

BASE COUNT 229 a 318 c 271 g 226 t

Query Match 75.4%; Score 865.8; DB 10; Length 1044;  
Best Local Similarity 90.3%; Pred. No. 6.1e-221;  
Matches 956; Conservative 0; Mismatches 82; Indels 21; Gaps 2;

QY 1 ATGCCAGAGCCGAGGTCCTTCTCTCTTCCCTGTCGCCCATCTCCATCTTTGTCTTC 60  
Db 1 ATGGCGAGTCCCCAGGTACCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60  
QY 61 GTGACTTCCACCATCATCCACCTCCAGCAGGATAGTGAAGCTCCACACCCCTGTACAG 120  
Db 61 GTGACTTCCACCATCATCCACCTCCAGCAGGATAGTGAAGCTCCACACCCCTGTACAG 120  
QY 121 AAGGAATTACCGATGACGACTCAATGCTCTCGGGAACACAGAACCCGACGATCGGA 180  
Db 121 AAGGAATTAC---AGCGGGTTCAAATGTCTCTCCACCAACCGGCAAGACAGCAT--- 173  
QY 181 CGGACAGCAGCAGCAGCATGGATGAGAGCTCGGGGGGATGTCAGATCAATTCATT 240  
Db 174 -----GCAGAGATGGCAAGCTCGAGGATATTCAGATCAATTCATT 219  
QY 241 GGCGGCTGGGAGACAGATGGGGAATAGCCACACTTTTCACATGGCCAGGATGAAC 300  
Db 220 GGCGGCTGGGAGACAGATGGGGAATAGCTATGCTATGTTGCACATGGCCAGGATGAAC 279  
QY 301 GGAGGGTTCGGTTCATTCCTCCGATCATGCAACAGCTTAGCGCCATCTTCAGGATC 360  
Db 280 GGTGGCTTCGCTTCATTCCTCCGATCATGCAACAGCTTAGCGCCATCTTCAGGATC 339  
QY 361 AGCTCCCGGTGTATACACGCGACAGCGCCAAAAGATCCATGGCAGATACCATCTC 420  
Db 340 AGTCTCCCGGTGTATACACGCGACAGCGAAGGATCCGCGGCAAGATACCATCTC 399  
QY 421 AACGACTGATGAGGAGGAGGCTTACCGCCACATCCGCGACACTTTGCGCTTCACGGGA 480  
Db 400 AACGACTGATGAGGAGGAGGCTTACCGCCACATCCCGGCGAGTATGTCGTTTCACGGGA 459  
QY 481 TACCGGTCTCTGGACCTTCTACACACCTCTGCGCCAGAGATCCGAGAGGATCAAC 540  
Db 460 TACCGGTCTCTGGACCTTCTACACACCTCTGCGCCAGAGATCCGAGAGGATCAAC 519  
QY 541 CTGCATGACCACTGCGGAGGAGGCGCGCTTCTGCTGGTGTCTGCGGGTGAATGGG 600  
Db 520 CTGCATGACCACTGCGGAGGAGGCGCGCTTCTGCTGGTGTCTGCGGGTGAATGGG 579  
QY 601 AGCCAGCGAGTACTTTTGGGTGTCCATGTGCGCGGAGGAGGACTATGTCATGTATG 660  
Db 580 AGCCAGCGAGTACTTTTGGGTGTCCATGTGCGCGGAGGAGGACTATGTCATGTATG 639  
QY 661 CTTATGTTGGAGGCGGTGCGGTGACCGGGGTTACCTGAAAGGCGCCCTGGACAG 720  
Db 640 CTTATGTTGGAGGCGGTGCGGTGACCGGGGTTACCTGAAAGGCGCCCTGGACAG 699  
QY 721 TTCGGGACGCTTATCATCTCCAGTCTTCTGCTGTACAGCAACGATGCGCTGGTGC 780  
Db 700 TTCGGGACGCTTATCATCTCCAGTCTTCTGCTGTACAGCAACGATGCGCTGGTGC 759  
QY 781 CGGAGAACATTAATGCTTCCGAGGAGGAGGCTGCTTCCGCGGCAATGTAATGAGGG 840  
Db 760 CGGAGAACATTAATGCTTCCGAGGAGGAGGCTGCTTCCGCGGCAATGTAATGAGGG 819

QY 841 TCGCCAGCCAGGAGGACITTCGGCTCTGCTCACCAGTGCACACACACATCATCATTTGGG 900  
Db 820 TCACAGCCAGGAGGACITTCGGCTCTGCTCACCAGTGCACACACACATCATCATTTGGG 879  
QY 901 ACCTTTGGGATTTGGGTGCTTACTTCCAGGTGGTGTATACCATCTACTTACTTACCACTAC 960  
Db 880 ACCTTTGGGATTTGGGTGCTTACTTCCAGGTGGTGTATACCATCTACTTACTTACCACTAC 939  
QY 961 ACCTTTGGGATTTGGGTGCTTACTTCCAAAGTCTTTAAGCAGGAGGAGGCTTCTTACCGGAA 1020  
Db 940 ACCTTTGGGATTTGGGTGCTTACTTCCAAAGTCTTTAAGCAGGAGGAGGCTTCTTCCCGAG 999  
QY 1021 TGGTGGGATTCCTGCGGATCTCTCCCATCTCTCTTAAG 1059  
Db 1000 TGGTGGGATTCCTGCGGATCTCTCCCATCTCTCTTAAG 1038

RESULT 9  
AB039114  
LOCUS  
DEFINITION  
Mus musculus FUT2 gene for GDP-L-fucose:beta-D-galactoside  
2-alpha-L-fucosyltransferase, partial cds, strain:C57BL/10SnJ.  
AB039114  
VERSION  
AB039114.1 GI:15822973  
KEYWORDS  
FUT2; GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase.  
SOURCE  
Mus musculus (sub-species:domesticus, strain:C57BL/10SnJ) DNA.  
ORGANISM  
Mus musculus  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1  
Liu,Y., Kitano,T., Koide,T., Shirolishi,T., Moriaki,K. and  
Saitou,N.  
Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes  
of Five Mus musculus subspecies  
Unpublished  
2 (bases 1 to 963)  
Liu,Y. and Saitou,N.  
Direct Submission  
TITLE  
Submitted (29-FEB-2000) Nariya Saitou, National Institute of  
Genetics, Laboratory of Evolutionary Genetics, 1111 Yata, Mishima,  
Shizuoka 411-8940, Japan (E-mail:nsaitou@genetics.nig.ac.jp,  
URL:http://sayer.lab.nig.ac.jp/, Tel:81-559-81-6790,  
Fax:81-559-81-6789)  
LOCATION/Qualifiers  
1..963  
/organism="Mus musculus"  
/strain="C57BL/10SnJ"  
/sub\_species="domesticus"  
/db\_xref="taxon:10090"  
/chromosome="7"  
/note="sequence used for primer design based on  
AccAF064792"

## FEATURES

source  
1..963  
/gene="FUT2"  
<1..>963  
/gene="FUT2"  
/EC\_number="2.4.1.69"  
/note="Sec2; Se"  
/codon\_start=1  
/product="GDP-L-fucose:beta-D-galactoside  
2-alpha-L-fucosyltransferase"  
/protein\_id="BAB68638.1"  
/db\_xref="GI:15822974"  
/translation="FLIFVFTSTIIHLQORIVKLTSEKELOAVQMSNSNARTDM  
QSQALQGLFTINSIGRLNOMGEYATLALANMGRALFIPESMINALAFIPRI  
VHSDTARPIWYHLNDWMEERYRHIPQYVYFTGYPCSTFVHRLPEILKEPTL  
HDHVRBEAFLGLRVNGSQPSTFVGVHVRDGYVHVPKVKWGVVADRGYLEKALD  
RFRARYSSPVVTSNMAWCRENINTSLGDDVFAVAGNIEGSPA KDFALLTQCCHNTIM  
TIGTGIWVAYLAGGDTIVLVNTLPDPSFLKIFKFAANFLPWW"

BASE COUNT  
214 a 285 c 257 g 207 t

Query Match  
69.8%; Score 802; DB 10; Length 963;

Best Local Similarity 90.2%; Pred. No. 7.6e-204;  
Matches 888; Conservative 0; Mismatches 75; Indels 21; Gaps 2;

QY 43 TTCCTCATCTTTGCTTGTGACTTCCACCATCATCCACTCCAGCAGCGAATAGTGAAG 102  
DB 1 TTCCTCATCTTTGCTTGTGACTTCCACCATCATCCACTCCAGCAGCGAATAGTGAAG 60  
QY 103 CTCACACCCCTCTCAGAGAAGAAATACCGATGAGAGACTCAATGATCCPCGGGAACACA 162  
DB 61 CTCACACCCCTCTCAGAGAAGAAATAC---AGGCGTTCAATGTCCTCACCAACCGG 117  
QY 163 GAAACCCAGATCGGACGGGACACCGAGCGATGGAATGGAGAGCTGCGGGCATG 222  
DB 118 GCAAGACAGACAT-----GCAGCAGAGTGCCTAAGCTGCAGGCGATA 159  
QY 223 TTCACGATCAATTCATTCGCGGCTGGGAACACAGATGGGGAATACCCACATCTTT 282  
DB 160 TTCACGATCAATTCATTCGCGGCTGGGAACACAGATGGGGAATATGCTACATGTTT 219  
QY 283 GCACTGGCCAGGATGAACGAGCGCTTGGCTTCATCCCGCATCCATGCACACGCTCTA 342  
DB 220 GCACITGGCCAGGATGAACGCTGGCTTGGCTTCATCCCTGAATCCATGCACACGCTCTA 279  
QY 343 GCGCCATCTTCAGATCAGCTCCCGGTTGTACACGCGACACGCGCCAAAGATCCCA 402  
DB 280 GCGCCATCTTCAGATCAGCTCCCGGTTGTACACGCGACACGCGCCAAAGATCCCG 339  
QY 403 TGGCAGATTAACATCTCAACCACTGGATGGAGGAGCGTTACCGGCACATTCGGGACAC 462  
DB 340 TGGCAGATTAACCACTCAACCACTGGATGGAGGAGCGTTACCGGCACATTCGGGACAC 399  
QY 463 TTTGTGGCTTCACGGGATACCGCTGCTCCGCTTCTTACACCACTGCGGCGCCACAG 522  
DB 400 TATGTGGCTTCACGGGATACCGCTGCTCCGCTTCTTACACCACTGCGGCGCCACAG 459  
QY 523 ATCCTGAGGAGTTCACTTCATCAGCAGCTGCGGAGGAGGCGCCAGCCCTTCCTGGT 582  
DB 460 ATCCTGAGGAGTTCACTTCATCAGCAGCTGCGGAGGAGGCGCCAGCCCTTCCTGGT 519  
QY 583 GTCTGCGGGTGAATGGGACCGACCGAGTACTTTTGTGGGTTCCACTGTCGCGCGAGG 542  
DB 520 GCGCTGCGGGTGAATGGGACCGACCGAGTACTTTTGTGGGTTCCACTGTCGCGCGAGG 579  
QY 643 GACTATGTCATGTCATCCCTAATGTGGAAGCGCTGGTGGCTGACCGGGTTACCTG 702  
DB 580 GACTATGTCATGTCATCCCAAGCTGTGGAAGCGCTGGTGGCTGACCGGGTTACCTA 639  
QY 703 GAAAGGCCCTGGATGTTCCGGCAGCGCTATTCATCTCCAGTCTTGTGTTTACAAGC 762  
DB 640 GAAAGGCCCTGGAGAGTTCCGGCAGCGCTATTCATCTCCAGTCTTGTGTTTACAAGC 699  
QY 763 AACGATATGGCTGTTGCGCGGAGAACATTAATGCTTCCGAGGAGAGCTGGTGTTCGCG 822  
DB 700 AACGATATGGCTGTTGCGCGGAGAACATTAATGCTTCCGAGGAGAGCTGGTGTTCGCG 759  
QY 823 GCAATGATATGAGGCTGCGCAGCAGGACTTCCGCTGCTCACCAGGATGCAACAC 882  
DB 760 GCAATGATATGAGGCTGCGCAGCAGGACTTCCGCTGCTCACCAGGATGCAACAC 819  
QY 883 ACCATCATGATATGAGGCTTGGGATTTGGGCTGCTCACCAGGATGCTGATACC 942  
DB 820 ACCATCATGATATGAGGCTTGGGATTTGGGCTGCTCACCAGGATGCTGATACC 879  
QY 943 ATCTACTTAGCACTACACCTTCCGATTTCCGTTCTCCCTCAAGTCTTTAAGCCAGAG 1002  
DB 880 ATCTACTTAGCACTACACCTTCCGATTTCCGTTCTCCCTCAAGTCTTTAAGCCAGCA 939  
QY 1003 GCAGCCTTCCCTACCGAATGGGTG 1026  
DB 940 GCAGCCTTCCCTACCGAATGGGTG 963

RESULT 10

AB039115  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
Source

AB039115 963 bp DNA linear ROD 02-OCT-2001  
Mus musculus FUT2 gene for GDP-L-fucose:beta-D-galactoside  
2-alpha-L-fucosyltransferase, partial cds, strain:BFM/2Msf.  
AB039115  
FUT2; GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase.  
Mus musculus (sub\_species:brevirostris, strain:BFM/2Msf) DNA.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Liu,Y., Kitano,T., Koide,T., Shirolishi,T., Moriwaki,K. and  
Saitou,N.  
Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes  
of Five Mus musculus subspecies  
Unpublished  
2 (bases 1 to 963)  
Liu,Y. and Saitou,N.  
Direct Submission  
Submitted (29-FEB-2000) Naruya Saitou, National Institute of  
Genetics, Laboratory of Evolutionary Genetics; lili.yata, Mishima,  
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,  
URL:http://sayer.lab.nig.ac.jp/, Tel:81-559-81-6790,  
Fax:81-559-81-6789),  
Location/Qualifiers

1..963  
/organism="Mus musculus"  
/strain="BFM/2Msf"  
/sub\_species="brevirostris"  
/db\_xref="taxon:10090"  
/chromosome="7"  
/note="sequence used for primer design based on  
Acc#AF064792"  
1..963  
/gene="FUT2"  
<1..>963  
/gene="FUT2"  
/EC\_number="2.4.1.69"  
/note="Sec2; Se"  
/codon\_start=1  
/product="GDP-L-fucose:beta-D-galactoside  
2-alpha-L-fucosyltransferase"  
/protein\_id="BAB68639.1"  
/db\_xref="GI:15822976"  
/translation="FLIFVFTVTIHLQORIVKLTSEKELQAVQMSNPNAATDM  
QSSALQGIPTINSIGRLNOMGEYATLALARMGRILAFIPESHNLALAPFRISLP  
VLHSTARRIPWQYHLNDMEERYHIPGOVRFYPCSWTFYHHLRPEIKFTLP  
HDHVEEAQAFILRLVNGSQPSTFVGVVRRGDYHVMKVKWGVADRGLKALD  
RFRARYSGPVFTVSGMAWCPENITSLGDVVRVFRAGNLEGGSPAKDFALLTQCNHTIM  
TIGTFGINAAVLAGSDTIVLANVTLPDSPFLKIFKPAAPLPEWM"  
BASE COUNT 214 a 285 c 257 g 207 t

Query Match 69.8%; Score 802; DB 10; Length 963;  
Best Local Similarity 90.2%; Pred. No. 7.6e-204;  
Matches 888; Conservative 0; Mismatches 75; Indels 21; Gaps 2;

QY 43 TTCCTCATCTTTGCTTGTGACTTCCACCATCATCCACTCCAGCAGCGAATAGTGAAG 102  
DB 1 TTCCTCATCTTTGCTTGTGACTTCCACCATCATCCACTCCAGCAGCGAATAGTGAAG 60  
QY 103 CTCACACCCCTCTCAGAGAAGAAATACCGATGAGAGACTCAATGATCCCTCCGCGGACACA 162  
DB 61 CTCACACCCCTCTCAGAGAAGAAATAC---AGGCGTTCAATGTCCTCACCAACCGG 117  
QY 163 GAAACCCAGATCGGACGGGACACAGCAGCATGGGAATGGAGAGCTGCGGGCATG 222  
DB 118 GCAAGACAGACAT-----GCAGCAGAGTGCCTAAGCTGCAGGCGATA 159  
QY 223 TTCACGATCAATTCATTCGCGGCTGGGGAACACAGATGGGGAATAGCCACATCTTT 282  
DB 160 TTCACGATCAATTCATTCGCGGCTGGGGAACACAGATGGGGAATATGCTACATGTTT 219

QY 283 GCACCTGGCCAGGATGAACGGCGCTTCGCTCATCCCGCATCATCACAACGCTCTA 342  
 Db 220 GCACCTGGCCAGGATGAACGGCGCTTCGCTCATCCCGCATCATCACAACGCTCTA 279  
 QY 343 GCGCCCATCTTCAGGATCAGGCTCCCGGTGTATACAGCGACGACGCGCAAAAGATCCCA 402  
 Db 280 GCGCCCATCTTCAGGATCAGCTCCCGGTGTATACAGCGACGACGCGCAAAAGATCCCG 339  
 QY 403 TGGGAGAAATACCATCTCAAGACATGAGTGGAGGCGGTACCGCCACATTCGCGGACAC 462  
 Db 340 TGGGAGAAATACCATCTCAAGACATGAGTGGAGGCGGTACCGCCACATTCGCGGCGAG 399  
 QY 463 TTGTGGCGCTTCACGGATACCGGTCTCCGACCTTCTACACACCTCGGCCAGAG 522  
 Db 400 TAATGGCGCTTCACGGATACCGGTCTCCGACCTTCTACACACCTCGGCCAGAG 459  
 QY 523 ATCTGAAGGAGTTACCTCTGATCAGGAGGCGGTGCGGAGGCGCGGCGGTTCGCTG 582  
 Db 460 ATCTGAAGGAGTTACCTCTGATCAGGAGGCGGTGCGGAGGCGCGGCGGTTCGCTG 519  
 QY 583 GGTCTGGCGGTGAATGGAGCGACCGAGTACTTTTGGGGTGTCCATGTGCGCGGAGGG 642  
 Db 520 GCGCTGGCGGTGAATGGAGCGACCGAGTACTTTTGGGGTGTCCATGTGCGCGGAGGG 579  
 QY 643 GACTATGTGCATGTCTCATCTAATCTGTGGAGGCGGTGGTGGCTGACCGGGTTCACCTG 702  
 Db 580 GACTATGTGCATGTCTCATCTAATCTGTGGAGGCGGTGGTGGCTGACCGGGTTCACCTA 639  
 QY 703 GAAAGGCGCTGGATGTATGTCGGGACGCTATTATCATCTCCAGTCTTGTGGTTACAAGC 762  
 Db 640 GAAAGGCGCTGGAGAGTTCGGGACGCTATTATCATCTCCAGTCTTGTGGTTACAAGC 699  
 QY 763 AACGTATGGCTGTGTCGGGAGAACATTAATGCTTCCGAGGAGAGTGTGTTCGCG 822  
 Db 700 AACGTATGGCTGTGTCGGGAGAACATTAATGCTTCCGAGGAGAGTGTGTTCGCG 759  
 QY 823 GCAATGATATGAGGGTTCGCCAGGAGTTCGGGCTGCTCACCAGTGCACACCAAC 882  
 Db 760 GCAATGATATGAGGGTTCGCCAGGAGTTCGGGCTGCTCACCAGTGCACACCAAC 819  
 QY 883 ACCATCATGACTATGGAGCTTTGGGATTTGGGTTGGCTGCTTACCTGGGAGTGGTATACC 942  
 Db 820 ACCATCATGACTATGGAGCTTTGGGATTTGGGTTGGCTGCTTACCTGGGAGTGGTATACC 879  
 QY 943 ATCTACTTAGCAACTACACCTTCGGATTCCTCGTTCTCCTCAAGTCTTTAAGCCAGAG 1002  
 Db 880 ATCTACTTAGCAACTACACCTTCGGATTCCTCGTTCTCCTCAAGTCTTTAAGCCAGCA 939  
 QY 1003 GCAGCCTTCCTACCGAATGGGTG 1026  
 Db 940 GCAGCCTTCCTACCGAATGGGTG 963  
 RESULT 11  
 AB039116  
 LOCUS  
 DEFINITION  
 2-alpha-L-fucosyltransferase, partial cds, strain: BLG2/Msf.  
 ACCESSION  
 AB039116  
 VERSION  
 AB039116.1  
 KEYWORDS  
 FUT2; GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase.  
 SOURCE  
 Mus musculus  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1  
 AUTHORS  
 Liu,Y., Kitano,T., Koide,T., Shiroishi,T., Moriaki,K. and  
 Saitou,N.  
 TITLE  
 Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes  
 of Five Mus musculus subspecies  
 JOURNAL  
 Unpublished  
 REFERENCE  
 2 (bases 1 to 963)

AUTHORS  
 TITLE  
 JOURNAL  
 Submitted (29-FEB-2000) Naruya Saitou, National Institute of  
 Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,  
 Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,  
 URL:http://sayer.lab.nig.ac.jp/, Tel:81-559-81-6790,  
 Fax:81-559-81-6789)  
 FEATURES  
 Location/Qualifiers  
 1..963  
 /organism="Mus musculus"  
 /strain="BLG2/Msf"  
 /sub\_species="musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="7"  
 /note="sequence used for primer design based on  
 Acc#AF064792"  
 gene  
 1..963  
 /gene="FUT2"  
 CDS  
 <1..>963  
 /gene="FUT2"  
 /EC\_number="2.4.1.69"  
 /note="Sec2; Se"  
 /codon\_start=1  
 /product="GDP-L-fucose:beta-D-galactoside  
 2-alpha-L-fucosyltransferase"  
 /protein\_id="BA68640.1"  
 /db\_xref="GI:15822978"  
 /translation="FLIEFVFTSTIIHQRIYKLTISEKELQAVQVSSNPAATDM  
 VLSKLOGITRINSIGRLNGMEYATIFALARMNGRLAFIPESHNALAFIFRISLP  
 QHSDATRIIPQWYHLNDWMEYRHPYGVYRTGYPCSWTFVHHRPIIKBFTL  
 HDHVREAAQFLRLGRVNGSOPSTFVGVHVRGDYVHKVPKVKWGVADRGLKALD  
 RPRARYSSPFEVVTSGMAMCRNINTSLGVDVFPAGNTEGSPADKFDALLTQCNHTIM  
 TIGTGIRWALYLAGDITLYLANYTLPDSPFLKIFKPAALPEW"  
 BASE COUNT 214 a 285 c 257 g 207 t  
 ORIGIN  
 Query Match 69.8%; Score 802; DB 10; Length 963;  
 Best Local Similarity 90.2%; Pred. No. 7.6e-204;  
 Matches 888; Conservative 0; Mismatches 75; Indels 21; Gaps 2;  
 QY 43 TTCTTCATCTTGTCTTCTGCTGACTTCCACCATCATCCACCTCCAGCAGCAATAGTGAG 102  
 Db 1 TTCTCTCATCTTGTCTTCTGCTGACTTCCACCATCATCCACCTCCAGCAGCAATAGTGAG 50  
 QY 103 CTCCAAACCTCTCAGAGAGGAATACCGATGAGCACTCAATATGCTCTCGGAAACACA 162  
 Db 61 CTCCAAACCTCTCAGAGAGGAATATAC---AGGCGTTCAATATGCTCTCCACCAACGG 117  
 QY 163 GAAAGCCACAGATCGCGGAGACAGCAGCAGCATGGGAATGGAGAGCTGCGGGGCATG 222  
 Db 118 GCAAGAACAGACAT-----GCAGCAGAGTGCCAAAGCTGCAGGGCATA 159  
 QY 223 TTCACGATCAATTCATTTGGCGGCTGGGACACAGATGGGGGAATACGCCACATCTTT 282  
 Db 160 TTCACGATCAATTCATTCGCGGCTGGGACACAGATGGGGGAATATGCTACATGTGTT 219  
 QY 283 GCACCTGGCCAGGATGAACGGCGGTTCGCTCATCCCGCATCATCACAACGCTCTA 342  
 Db 220 GCACCTGGCCAGGATGAACGGCGGTTCGCTCATCCCGCATCATCACAACGCTCTA 279  
 QY 343 GCGCCCATCTTCAGGATCAGCCTCCCGGTGTATACAGCGACGACGCGCAAAAGATCCCA 402  
 Db 280 GCGCCCATCTTCAGGATCAGCTCCCGGTGTATACAGCGACGACGCGCAAAAGATCCCG 339  
 QY 403 TGGGAGAAATACCATCTCAAGACATGAGTGGAGGCGGTACCGCCACATTCGCGGACAC 462  
 Db 340 TGGGAGAAATACCATCTCAAGACATGAGTGGAGGCGGTACCGCCACATTCGCGGCGAG 399  
 QY 463 TTGTGGCGCTTCACGGATACCGGTCTCCGACCTTCTACACACCTCGGCCAGAG 522  
 Db 400 TAATGGCGCTTCACGGATACCGGTCTCCGACCTTCTACACACCTCGGCCAGAG 459  
 QY 523 ATCTGAAGGAGTTACCTCTGATCAGGAGGCGGTGCGGAGGCGCGGCGGTTCGCTG 582



```

Db      460 ATCTTGAAGAGTACCTGACGACCATGTCGTGAGGAGCCAGGCTTTCCTGCGT 519
Qy      583 GGTCTGCGGGTGAATGGGACGACGAGTACTTTTGTGGTGTCCATGTGCGCCAGGG 642
Db      520 GGCTCGCGGTGAATGGGACGACGAGTACTTTTGTGGTGTCCATGTGCGCCAGGG 579
Qy      643 GACTATGTCATGTCATGCTAATGTGTGGAAGGCGTGTGCTGACCGGGTTACCTG 702
Db      580 GACTATGTCATGTCATGCTAATGTGTGGAAGGCGTGTGCTGACCGGGTTACCTG 639
Qy      703 GAAAGGCGCTGATATGTCGCGGACGCTATTCATTCCTGAGTCTTCTGTTTAAAGC 762
Db      640 GAAAGGCGCTGATATGTCGCGGACGCTATTCATTCCTGAGTCTTCTGTTTAAAGC 699
Qy      763 AACGATATGCTGCTGCTGCGGAGAACATTAATGCTCCGAGGAGACGTGTGTCGG 822
Db      700 AACGATATGCTGCTGCTGCGGAGAACATTAATGCTCCGAGGAGACGTGTGTCGG 759
Qy      823 GCATATGCTGCTGCTGCGGAGAACATTAATGCTCCGAGGAGACGTGTGTCGG 882
Db      760 GCATATGCTGCTGCTGCGGAGAACATTAATGCTCCGAGGAGACGTGTGTCGG 819
Qy      883 ACCATCATGATGAGGAGCTTGGGATTTGGGCTGCTTACCTGCGAGGAGTGTGATCC 942
Db      820 ACCATCATGATGAGGAGCTTGGGATTTGGGCTGCTTACCTGCGAGGAGTGTGATCC 879
Qy      943 ATCTACTTAGCCAACTACACCTTCCGATTTCCGATTTCCGATTTCCGATTTCCGATTT 1002
Db      880 ATCTACTTAGCCAACTACACCTTCCGATTTCCGATTTCCGATTTCCGATTTCCGATTT 939
Qy      1003 GCAGCTTCCCTACCGATGGGTG 1026
Db      940 GCAGCTTCCCTACCGATGGGTG 963

RESULT 12
LOCUS   AB039117
DEFINITION Mus musculus FUN2 gene for GDP-L-fucose:beta-D-galactoside
ACCESSION AB039117
VERSION 1
KEYWORDS FUN2; GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Liu,Y., Kitano,T., Koide,T., Shiroishi,T., Moriwaki,K. and Saitou,N.
TITLE Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes
JOURNAL of Five Mus musculus subspecies
REFERENCE 2
AUTHORS Liu,Y. and Saitou,N.
TITLE Direct Submission
JOURNAL Submitted (29-FEB-2000) Naruya Saitou, National Institute of Genetics, Laboratory of Evolutionary Genetics, 1111 Yata, Mishima, Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp, URL:http://sayer.lab.nig.ac.jp/, Tel:81-559-81-6790, Fax:81-559-81-6789)
FEATURES             Location/Qualifiers
     source            1..963
                        /organism="Mus musculus"
                        /strain="CAST/Ei"
     sub_species="castaneus"
     db_xref="taxon:10090"
     chromosome="7"
     note="sequence used for primer design based on Acc#AF064792"
     1..963
     /gene="FUN2"

```

## CDS

```

<1..>963
/gene="FUN2"
/EC_number="2.4.1.69"
/note="Sec2; Se"
/codon_start=1
/product="GDP-L-fucose:beta-D-galactoside
2-alpha-L-fucosyltransferase"
/protein_id="BAB6841.1"
/db_xref="GI:15822980"
/translation="FLIFVFTSTIIHQORIVKIQTLSEKELQAVOMSSPNAARDTM
QSAKLQFTINSIRGLNQMEYATLALARMNGRLAIPESMNAALAFIRISLP
VLHSTARRIPWNLHNDWERYRHIPQXVRFYPCSWTFYHLPEILKEETL
HDBVEAQAFLRLVNGSQSTFVGHVVRGDDVYHVKVGVKGVADRGYLEKALD
RRARYSSPFVFTVSNMANKRENINTSLGDVVFAGNGIEGSPAKFDALTGCTNHTIM
TIGTGTWAAIYLAGDTIYLANITLPSFLKIFKPAAPLPEWM"
BASE COUNT 214 a 285 c 257 g 207 t
ORIGIN
Query Match 69.8%; Score 802; DB 10; Length 963;
Best Local Similarity 90.2%; Pred. No. 7.6e-204;
Matches 888; Conservative 0; Mismatches 75; Indels 21; Gaps 2;
Qy 43 TTCCATCATTTTGTCTCTGCTGACTTCCACCATCATCCACTCCAGAGGGAATAGTAGG 102
Db 1 TTCCATCATTTTGTCTCTGCTGACTTCCACCATCATCCACTCCAGAGGGAATAGTAGG 60
Qy 103 CTCCACCCCTCTCAGAGAGGAATACCGATGACGACTCAATATGCTCTGGGGAACA 162
Db 61 CTCCACCCCTCTCAGAGAGGAATATAC---AGGCGGTTCAATGTCTCTCAACAACGG 117
Qy 163 GAAAGCCAGATGCGACGGGACAGCGAGCATGGGAATGGAGAGCTCGGGGCGATG 222
Db 118 GCAAGAACAGACAT-----GCAGCAGATGCCAAGCTGCAGGGCATA 159
Qy 223 TTCAGATCAATTCATGGCGGCTGGGGAACACAGATGGGCGAATACGCCACACTTT 282
Db 160 TTCAGATCAATTCATGGCGGCTGGGGAACACAGATGGGCGAATATCTACATGTT 219
Qy 283 GCATGGCAGATCAAGGAGGCTTGGTTCATCCCGGATCCCGCATCCATGACACGCT 342
Db 220 GCATGGCAGATCAAGGAGGCTTGGTTCATCCCGGATCCCGCATCCATGACACGCT 279
Qy 343 GGGCCATCTTCAGGATCAGCTCCCGGTGTGTACACAGCAGCAGCAGCAAAAGATCCCA 402
Db 280 GGGCCATCTTCAGGATCAGCTCCCGGTGTGTACACAGCAGCAGCAGCAAGATCCCG 339
Qy 403 TGGCAGATTAACCATCTCAAGACTGGATGGAGAGCGTTACCGCACATTCGCGGAGAC 462
Db 340 TGGCAGATTAACCATCTCAAGACTGGATGGAGAGCGTTACCGCACATTCGCGGAGAC 399
Qy 463 TTTGTGCGCTTACGGGATACCGGTCTCTGACCTTCTACCCACCTGCGCCAGAG 522
Db 400 TATGTGCGCTTACGGGATACCGGTCTCTGACCTTCTACCCACCTGCGCCAGAG 459
Qy 523 ATCTGAAGAGTTCACCTGATGACCATGCTGCGGAGGAGCGCCAGCGCTTCCTGCGT 582
Db 460 ATCTGAAGAGTTCACCTGATGACCATGCTGCGGAGGAGCGCCAGCGCTTCCTGCGT 519
Qy 583 GCTGTGCGGCTGAATGGAGCCAGCGGAGTACTTTTGTGGTGTCCATGTGCGCGAGGG 642
Db 520 GCGTGTGCGGCTGAATGGAGCCAGCGGAGTACTTTTGTGGTGTCCATGTGCGCGAGGG 579
Qy 643 GACTATGTCATGTCATGCTAATGTGTGGAAGGCGTGTGCTGACCGGGTTACCTG 702
Db 580 GACTATGTCATGTCATGCTAATGTGTGGAAGGCGTGTGCTGACCGGGTTACCTG 639
Qy 703 GAAAGGCGCTGATATGTCGCGGACGCTATTCATTCCTGAGTCTTCTGTTTAAAGC 762
Db 640 GAAAGGCGCTGATATGTCGCGGACGCTATTCATTCCTGAGTCTTCTGTTTAAAGC 699
Qy 763 AACGATATGCTGCTGCTGCGGAGAACATTAATGCTTCCGAGGAGAGACGTGTGTCGG 822
Db 700 AACGATATGCTGCTGCTGCGGAGAACATTAATGCTTCCGAGGAGAGACGTGTGTCGG 759

```



```

RESULT 14
AB039119
LOCUS
DEFINITION
Mus musculus FUT2 gene for GDP-L-fucose:beta-D-galactoside
2-alpha-1-fucosyltransferase, partial cds, strain:pgn2.
ACCESSION
AB039119.1
VERSION
GI:15822983
KEYWORDS
FUT2; GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase.
SOURCE
Mus musculus (sub_species:domesticus, strain:pgn2) DNA.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Liu,Y., Kitano,T., Koide,T., Shirolishi,T., Moriwaki,K. and
Saitou,N.
Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes
of Five Mus musculus subspecies
JOURNAL
Unpublished
2 (bases 1 to 963)
Liu,Y. and Saitou,N.
Direct Submission
Submitted (29-FEB-2000)
Naruya Saitou, National Institute of
Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
URL:http://sayer.lab.nig.ac.jp/, Tel:81-559-81-6790,
Fax:81-559-81-6789)
FEATURES
Location/Qualifiers
1..963
/organism="Mus musculus"
/strain="pgn2"
/sub_species="domesticus"
/db_xref="taxon:10090"
/chromosome="7"
/note="sequence used for primer design based on
acc#AF064792"
1..963
/ gene="FUT2"
/ CDS
<1..>963
/ gene="FUT2"
/ EC_number="2.4.1.59"
/note="Sec2; Se"
/codon_start=1
/product="GDP-L-fucose:beta-D-galactoside
2-alpha-1-fucosyltransferase"
/protein_id="BAB68643.1"
/db_xref="GI:15822984"
/translation="FLIFVFTSTIIHQRIKVLQSLSEKELQVQSSNNAATDM
QOSKLGIFITNSIGRUGNMGAYATLFLARMNGRLAIFPESMHALAIFRISLP
VLSDTARRIPQWNYHLNDMEERYRHIPGOYVRFYPCSWIFVHHLRPEILKEFTL
HDHVEEAQAFRLGLRVNGSQSPSTFVGVHVRGQYVHVMPKWKGVADRQYLEKALD
RFRARYSSPVFVTSNGMAWCNRENTSLGVDVPAAGNIEGSPAKDFALLTQCNIHTM
TIGTFGIWAAYLAGSDIYLYANLYLPDSPEFLKIKPAAFLPEWM"
BASE COUNT
214 a 285 c 257 g 207 t
ORIGIN
Query Match 69.8%; Score 802; DB 10; Length 963;
Best Local Similarity 90.2%; Pred. No. 7.6e-204;
Matches 888; Conservative 0; Mismatches 75; Indels 21; Gaps 2;
QY 43 TTCCTCATCTTGTGCTTGGTACCTCCACCATCATCCACCTCCACGACGATAGTGAAG 102
Db [|||||]
1 TTCCCTCATCTTGTGCTTGGTACCTCCACCATCATCCACCTCCACGACGATAGTGAAG 60
QY 103 CTCCAACCCCTGTCAGAGAAGAAATACCGACTGACGACTCAAAATGCTCTCGGAACACA 162
Db [|||||]
61 CTCCAACCCCTGTCAGAGAAGAAATAC---AGCGGTTCAAATGCTCTCACCACCAACCG 117
QY 163 GAAGCCGACAGATGCGGACGGGACGACGACGATGGGAATGGAGAGCTCGCGGGCATG 222
Db [|||||]
118 GCAAGACAGACAT-----GCAGCAGAGTGCACAGCTCAGGCGCAT 159
QY 223 TTCAGTCATTTCCATGGCCGCTGGGGACACAGATGGCGGATAGCCACACTCTTT 282
Db [|||||]

```

```

Db 160 TTCACGTCATATCCATCGCGCGCTGGGGAACAGATGGCGGAATATGCTACATTTT 219
QY 283 GCATGCCCCAGGATGAACGACGCGCTTGCCTTCATCCCCGATCCATGCAACAGCTCTA 342
Db [|||||]
220 GCATGCCCCAGGATGAACGCGTGGCTTCATCCCTGAATCCATGCACAGCTCTA 279
QY 343 GCGCCCATCTTCAGGATCAGCCTCCCGGTGTTACACAGCAGACAGGCGCAAAAGATCCCA 402
Db [|||||]
280 GCGCCCATCTTCAGGATCAGCTCCCGGTGTTACACAGCAGACAGGCGCAAGAGATCCCG 339
QY 403 TGCAGAAATACCATCTCAACGACTGGATGAGGAGGTTACCGCCACATTCGGGACAC 462
Db [|||||]
340 TGCAGAAATACCATCTCAACGACTGGATGAGGAGGTTACCGCCACATTCGGGCGAG 399
QY 463 TTTGTGCGCTTCACGGGATACCCGCTCTCTCGACCTTCTACACCACTTCGCGCCAGAG 522
Db [|||||]
400 TATGTGCGTTTCACGGGATACCCGCTCTCTCGACCTTCTACACCACTTCGCGCCAGAG 459
QY 523 ATCCTGAAGAGTTTACCCCTGCGATGACGAGTGGCGGAGGAGGCGGCGCTTCCTGCGGT 582
Db [|||||]
460 ATCCTGAAGAGTTTACCCCTGCGATGACGAGTGGCGGAGGAGGCGGCGCTTCCTGCGGT 519
QY 583 GGTCTGCGGTGAATGGGAGCGGAGTACTTTTGGGTGCTCATGTGCGCGGAGG 642
Db [|||||]
520 GCGCTGCGGTGAATGGGAGCGGAGTACTTTTGGGTGCTCATGTGCGCGGAGG 579
QY 643 GACTATGTCATGTCATGCTTAATGTGGAAGGCGTGTGCTGACCGGGGTTCCTGTG 702
Db [|||||]
580 GACTATGTCATGTCATGCTTAATGTGGAAGGCGTGTGCTGACCGGGGTTCCTGTG 639
QY 703 GAAAGGCGCTGGATATGTTCCGGGACGCTATTCATCTCCAGTCTTCGTGGTACAAGC 762
Db [|||||]
640 GAAAGGCGCTGGAGAGTTCCGGGACGCTATTCATCTCCAGTCTTCGTGGTACAAGC 699
QY 763 AACGCTATGCGCTGTGCGGGAGAACATTAATGCTTCCGAGGAGGAGTGTTCGCG 822
Db [|||||]
700 AACGCTATGCGCTGTGCGGGAGAACATTAATGCTTCCGAGGAGGAGTGTTCGCG 759
QY 823 GGCATGCTATGAGGGGTGCGGAGGAGTGTGCGCTGCTCACCAGTGTCAACACAC 882
Db [|||||]
760 GGCATGCTATGAGGGGTGCGGAGGAGTGTGCGCTGCTCACCAGTGTCAACACAC 819
QY 883 ACCATCATGACTATGCGGACCTTGGGATTTGGGCTGCTACCTGCGAGGTTGTAIAC 942
Db [|||||]
820 ACCATCATGACTATGCGGACCTTGGGATTTGGGCTGCTACCTGCGAGGTTGTAIAC 879
QY 943 ATCTACTAGCCAACTACACCTTCCGATTTCCGATTTCCGATTTCCGATTTCCGATTT 1002
Db [|||||]
880 ATCTACTAGCCAACTACACCTTCCGATTTCCGATTTCCGATTTCCGATTTCCGATTT 939
QY 1003 GCAGGCTTCCCTACCGAATGGGTG 1026
Db [|||||]
940 GCAGGCTTCCCTACCGAATGGGTG 963

RESULT 15
AB039120
LOCUS
DEFINITION
Mus musculus FUT2 gene for GDP-L-fucose:beta-D-galactoside
2-alpha-1-fucosyltransferase, partial cds, strain:SWN/Msf.
ACCESSION
AB039120.1
VERSION
GI:15822985
KEYWORDS
FUT2; GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase.
SOURCE
Mus musculus (sub_species:molossinus, strain:SWN/Msf) DNA.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Liu,Y., Kitano,T., Koide,T., Shirolishi,T., Moriwaki,K. and
Saitou,N.
Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes
of Five Mus musculus subspecies
JOURNAL
Unpublished

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 09:17:27 ; Search time 298.004 Seconds  
(without alignments)  
8682.922 Million cell updates/sec

Title: US-10-040-863-7

Perfect score: 1149

Sequence: 1 atggccagccacgattcc.....gaagagccttctgatgggaa 1149

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1149	100.0	1149 21	Rat hepatoma H35 c
2	1149	100.0	1149 24	Rat hepatoma H35 c
3	1068	93.0	1068 21	Rat hepatoma H35 c
4	1068	93.0	1068 24	Rat hepatoma H35 c
5	630.2	54.8	1043 19	Porcine secretor t
6	630.2	54.8	1043 19	Pig secretor encod
7	613.8	53.4	2115 19	Human Sec2 coding
8	612.6	53.3	3088 24	Lung cancer relate
9	612.6	53.3	3088 24	Lung cancer relate

10	612.6	53.3	3088	24	ABL70026
11	423.8	36.9	1098	19	AAV21640
12	423.8	36.9	1269	20	AAI15872
13	423.8	36.9	1269	20	AAI15872
14	423.8	36.9	1269	21	AAI15872
15	423.8	36.9	1269	21	AAI15872
16	405	35.2	1155	16	AAQ98461
17	405	35.2	1155	16	AAQ98461
18	405	35.2	1155	16	AAQ98461
19	405	35.2	1174	17	AAQ98461
20	405	35.2	1174	17	AAQ98461
21	405	35.2	1174	17	AAQ98461
22	405	35.2	1174	17	AAQ98461
23	405	35.2	1174	17	AAQ98461
24	405	35.2	1174	17	AAQ98461
25	405	35.2	1174	17	AAQ98461
26	403.4	35.1	8174	18	AAQ98461
27	304.8	26.5	1071	20	AAQ98461
28	141.4	12.3	473	21	AAQ98461
29	72.8	6.3	100	21	AAQ98461
30	71.2	6.2	100	21	AAQ98461
31	71.2	6.2	100	21	AAQ98461
32	69.6	6.1	100	21	AAQ98461
33	63.2	5.5	100	21	AAQ98461
34	63.2	5.5	100	21	AAQ98461
35	55.8	4.9	100	21	AAQ98461
36	49.8	4.3	79	21	AAQ98461
37	44	3.8	495	22	AAQ98461
38	42.8	3.7	4059	22	AAQ98461
39	42.8	3.7	5236	22	AAQ98461
40	42.8	3.7	5236	22	AAQ98461
41	40.4	3.5	2297	22	AAQ98461
42	40.4	3.5	2612	22	AAQ98461
43	40.4	3.5	2953	22	AAQ98461
44	40.4	3.5	3044	22	AAQ98461
45	40.4	3.5	3044	22	AAQ98461

#### ALIGNMENTS

RESULT 1  
AAC67965  
ID AAC67965 standard; cDNA; 1149 BP.  
XX AAC67965;  
AC  
XX  
19-FEB-2001 (first entry)  
XX  
Rat hepatoma H35 cell alpha1-2fucosyltransferase cDNA.  
DE  
XX  
Rat; alpha1-2fucosyltransferase; cytotstatic; neuroprotective;  
KW nontropic; gene therapy; Fucalalpha1-2GALbeta1-3GALNAc; immunotherapy;  
KW immunosuppression; cancer; neurological disease;  
KW small cell lung carcinoma; ss.  
XX  
OS Rattus norvegicus.  
XX  
PN WO200064464-A1.  
XX  
PD 02-NOV-2000.  
XX  
PF 23-APR-1999; 99WO-US07384.  
XX  
23-APR-1999; 99WO-US07384.  
(PACI-) PACIFIC NORTHWEST CANCER FOUND.  
PI Holmes EH, Sherwood AL;  
XX  
WPI: 2000-687262/57.  
DR P-PSDB; AAB36104.  
XX



DR WPI; 2002-121132/16.  
DR P-PSDB; AAB16622.  
XX Rat hepatoma H35 cell alpha1-2fucosyltransferase, useful for producing  
PT GM1-specific alpha1-2fucosyltransferase enzyme by recombinant  
PT techniques and for detecting oncogenic transformation of test tissues -  
XX  
XX Claim 1; Fig 5; 4lpp; English.  
XX The invention relates to rat GM1-specific alpha1-2fucosyltransferase  
CC (alpha1-2Fuct) enzyme and its corresponding nucleic acid. This nucleic  
CC acid is specific for a carbohydrate moiety found in ganglioside GM1,  
CC a terminal galactose beta1-3N-acetylgalactosamine (Galbeta1-3GalNAc)  
CC saccharide. Alpha1-2Fuct DNA is useful for producing rat alpha1-2Fuct  
CC protein by recombinant techniques. Alpha1-2Fuct DNA is useful for the  
CC preparative synthesis of fucosyl containing glycolipids, glycoproteins,  
CC glycolipoproteins and oligosaccharide, and for preparing fucosyl-GM1.  
CC Alpha1-2Fuct DNA is useful for detecting oncogenic transformation which  
CC involves assaying for changes in expression of alpha1-2 Fuct. Since  
CC alpha1-2Fuct is activated in cell transformation, antisense sequences  
CC derived from alpha1-2Fuct DNA are useful for inhibiting, suppressing  
CC or treating cancer. Alpha1-2Fuct DNA is useful in gene therapy and  
CC antisense therapy. The present sequence is rat hepatoma H35 Cell  
CC alpha1-2Fuct DNA.  
XX  
XX Sequence 1149 BP; 247 A; 341 C; 311 G; 250 T; 0 other;  
Query Match 100.0%; Score 1149; DB 24; Length 1149;  
Best Local Similarity 100.0%; Pred. No. 1.8e-310;  
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCAGGCGCCAGGTCCTTTCTCTTCCTTCCTGGGCCACTTCCTCATCTTTGCTTC 60  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1 ATGGCAGGCGCCAGGTCCTTTCTCTTCCTTCCTGGGCCACTTCCTCATCTTTGCTTC 60  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 61 GTGACTTCACCATCATCCACTCCAGCAGCAAGTAGTGAAGCTCCACCCCTGTACAGAG 120  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 61 GTGACTTCACCATCATCCACTCCAGCAGCAAGTAGTGAAGCTCCACCCCTGTACAGAG 120  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 121 AAGGAATACCGATGACGACTCAATGTCTCTCGGAAACACAGAAAGCCAGAGATGCGA 180  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 121 AAGGAATACCGATGACGACTCAATGTCTCTCGGAAACACAGAAAGCCAGAGATGCGA 180  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 181 CGGACAGCGAGCAGATGGGAATGGAGAGCTGCGGGGCATGTTCAAGATCAATTCATTT 240  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 181 CGGACAGCGAGCAGATGGGAATGGAGAGCTGCGGGGCATGTTCAAGATCAATTCATTT 240  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 241 GCGCGCTGGGGAACAGATGGGCGGANTACGCCACACTCTTTGCACTGCCAGGATGAAC 300  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 241 GCGCGCTGGGGAACAGATGGGCGGANTACGCCACACTCTTTGCACTGCCAGGATGAAC 300  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 301 GGACGGCTTGCTTCATCCCGCATCCATGCAACAGCTCTAGCGGCCATCTTCAGGATC 360  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 301 GGACGGCTTGCTTCATCCCGCATCCATGCAACAGCTCTAGCGGCCATCTTCAGGATC 360  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 361 AGCCTCCCGGTTTACAGAGCAACCGGCCAAAGATCCCATGGCAGAAATACCATCTC 420  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 361 AGCCTCCCGGTTTACAGAGCAACCGGCCAAAGATCCCATGGCAGAAATACCATCTC 420  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 421 AACGACTGGATGGAGAGCGGTTACGCCACATCTTCGGGACACTTTGTGGCTTCAGGGA 480  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 421 AACGACTGGATGGAGAGCGGTTACGCCACATCTTCGGGACACTTTGTGGCTTCAGGGA 480  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 481 TACCGGTGCTCTGACCTTCTACCAACACCTTCGCCAGAGATCCTGAAGAGTTTCAAC 540  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 481 TACCGGTGCTCTGACCTTCTACCAACACCTTCGCCAGAGATCCTGAAGAGTTTCAAC 540  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 541 CAGCATGACCACTGCGGAGAGAGCCAGGCTTCCTCGGTGCTGCGGAGTGAATGGG 600  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 541 CAGCATGACCACTGCGGAGAGAGCCAGGCTTCCTCGGTGCTGCGGAGTGAATGGG 600  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 601 ACCGACCGAGTACTTTGTGGGTTCATGTGCCGAGGGGACTATGTGCATCTCATG 660  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 601 AGCCAGCCGAGTACTTTTCTGGGTGCTCCATGTGCGCCGAGGGGACTATGTGCATGTCATG 660  
QY 661 CTAATCTGTGAAGGCGGTGCTGCTGACCGGGGTACCTGGAAGAGCCCTTGGATATG 720  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 661 CTAATCTGTGAAGGCGGTGCTGCTGACCGGGGTACCTGGAAGAGCCCTTGGATATG 720  
QY 721 TTTCCGGGACAGCTATTCATCTCCAGTCTTCTGGGTGTACAGCAAGCGTATGGCTGCTGC 780  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 721 TTTCCGGGACAGCTATTCATCTCCAGTCTTCTGGGTGTACAGCAAGCGTATGGCTGCTGC 780  
QY 781 CGGAGAAACATTAATGCTTCCCGAGAGAGCGTGTGTTCCCGGGCAATGATTTGAGGG 840  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 781 CGGAGAAACATTAATGCTTCCCGAGAGAGCGTGTGTTCCCGGGCAATGATTTGAGGG 840  
QY 841 TCGCCAGCAAGGACTTCCGCTGCTCACCAGTGCACCCAGTGCACCCAGTGCATGATTTGG 900  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 841 TCGCCAGCAAGGACTTCCGCTGCTCACCAGTGCACCCAGTGCACCCAGTGCATGATTTGG 900  
QY 901 ACCTTTGGGATTTGGCTGCTACCTGCGAGGTGGTATACCATCTACTTAGCCAACTAC 960  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 901 ACCTTTGGGATTTGGCTGCTACCTGCGAGGTGGTATACCATCTACTTAGCCAACTAC 960  
QY 961 ACCCTTCGGGATTTCCGTTCTCTCAAAGTCTTTAAGCCAGAGGAGCCCTTCCTACCCGAA 1020  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 961 ACCCTTCGGGATTTCCGTTCTCTCAAAGTCTTTAAGCCAGAGGAGCCCTTCCTACCCGAA 1020  
QY 1021 TGGGTGGGATCCCTGCGGATCTGTCCTCCACCTTAAAGCATTAACACAGCCCTCTCT 1080  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 1021 TGGGTGGGATCCCTGCGGATCTGTCCTCCACCTTAAAGCATTAACACAGCCCTCTCT 1080  
QY 1081 CGTCCCTCACTTCCACCTCAAGGCAAGGAGTCACTGTTAGTCTCGAGAGAGAGCTTC 1140  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 1081 CGTCCCTCACTTCCACCTCAAGGCAAGGAGTCACTGTTAGTCTCGAGAGAGAGCTTC 1140  
QY 1141 TGATGGGAA 1149  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 1141 TGATGGGAA 1149

RESULT 3  
AAC67966  
ID AAC67966 standard; cDNA; 1068 BP.  
AC AAC67966;  
XX  
XX 19-FEB-2001 (first entry)  
DE Rat hepatoma H35 cell alpha1-2fucosyltransferase catalytic domain cDNA.  
XX  
XX Rat; alpha1-2fucosyltransferase; cytostatic; neuroprotective;  
KW nontropic; gene therapy; Fucalalpha1-2Galbeta1-3GalNAc; immunotherapy;  
KW immunosuppression; cancer; neurological disease;  
KW small cell lung carcinoma; ss.  
XX  
OS Rattus norvegicus.  
XX  
XX WO200064464-A1.  
PN  
XX  
PD 02-NOV-2000.  
XX  
XX 23-APR-1999; 99WO-US07384.  
PF  
XX  
XX 23-APR-1999; 99WO-US07384.  
PR  
XX  
XX (PACI-) PACIFIC NORTHWEST CANCER FOUND.  
PA  
XX  
PI Holmes EH, Sherwood AL;  
XX  
XX WPI: 2000-687262/67.  
DR P-PSDB; AAB36105.  
XX  
XX New rat ganglioside GM1-specific alpha1-2fucosyltransferase, useful for  
PT preparation of fucosyl GM1 which is useful as a nutritional composition

or immunotherapeutic for cancer and neurological diseases -

Claim 11; Fig 3A; 91pp; English.

The present sequence is given in a specification relating to a rat ganglioside GM1-specific alpha1-2fucosyltransferase protein. The protein or its cellular fraction is useful for synthesis of a molecule comprising Fucal-2Gal-beta1-3GalNAc, a glycolipid, glycoprotein, glycolipoprotein or a free oligosaccharide comprising Fucal-2Gal-beta1-3GalNAc. The method involves contacting alpha1-2fucosyltransferase with GDP-fucose and a molecule or glycolipid, glycoprotein, glycolipoprotein or oligosaccharide having a terminal Gal-beta1-3GalNAc group. It is also useful for synthesis of fucosyl-GM1 by contacting the protein with GDP-fucose and ganglioside GM1. The obtained glycoproteins, glycolipoproteins, glycolipids and oligosaccharides are useful as nutritional compositions and fucosyl-GM1 is useful for inducing an immunotherapeutic or immunosuppressive action against cancer, neurological disease or small cell lung carcinoma.

Sequence 1068 BP; 237 A; 308 C; 300 G; 223 T; 0 other;

Query Match 93.0%; Score 1068; DB 21; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 7,3e-288;  
Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 82 CTCGACGAGCGAATGTAAGCTCCACCCCTGTCCAGAGAAGATACCGATGACGACT 141
Db 1 CTCGACGAGCGAATGTAAGCTCCACCCCTGTCCAGAGAAGATACCGATGACGACT 60

QY 142 CAATGTCTCTCGGGAACACACAGAACGCCAGATGGACGGGACAGGACGATGGG 201
Db 61 CAATGTCTCTCGGGAACACACAGAACGCCAGATGGACGGGACAGGACGATGGG 120

QY 202 AATGAGAGCTGCGGGGCTATTCACGATCAATTCATGCGCGCTGGGGAACACAGT 261
Db 121 AATGAGAGCTGCGGGGCTATTCACGATCAATTCATGCGCGCTGGGGAACACAGT 180

QY 262 GGCAGATACGCCACACCTCTTGCACTGCCAGGATGAACGAGCGGCTGCGTTCATCCCC 321
Db 181 GGCAGATACGCCACACCTCTTGCACTGCCAGGATGAACGAGCGGCTGCGTTCATCCCC 240

QY 322 GCATCCATGACACAGCTCTAGCGCCCTATCTCAGGATCAGCCTCCGGTGTACAGC 381
Db 241 GCATCCATGACACAGCTCTAGCGCCCTATCTCAGGATCAGCCTCCGGTGTACAGC 300

QY 382 GACAGCGCCAAAAGATCCCATGCGAGAAITACCATCTCAACGACTGGATGGAGAGCGT 441
Db 301 GACAGCGCCAAAAGATCCCATGCGAGAAITACCATCTCAACGACTGGATGGAGAGCGT 360

QY 442 TACGCCACATTCGGGACACTTTGTGGCTTACGGGATACCGGTCTCTCTGGACCTTC 501
Db 361 TACGCCACATTCGGGACACTTTGTGGCTTACGGGATACCGGTCTCTCTGGACCTTC 420

QY 502 TACCACCACTGCGCCAGAGATCTCAAGAGATTCACCTGCATGACCAGCTCGGGAG 561
Db 421 TACCACCACTGCGCCAGAGATCTCAAGAGATTCACCTGCATGACCAGCTCGGGAG 480

QY 562 GAGCCCGAGCGCTTCCTCGCTGCTCGGTGAATGGAGCCAGCGAGTACTTTGTG 621
Db 481 GAGCCCGAGCGCTTCCTCGCTGCTCGGTGAATGGAGCCAGCGAGTACTTTGTG 540

QY 622 GGTGTCCATGTGCGCCAGGGGACTATGTGATGTCATGTCCTAATGTGTGAAGGGCGTG 681
Db 541 GGTGTCCATGTGCGCCAGGGGACTATGTGATGTCATGTCCTAATGTGTGAAGGGCGTG 600

QY 682 GTGCTGACCGGGGTACCTCGGAAGGCCCTGGATATGTCGGGACGCTATTCATCT 741
Db 601 GTGCTGACCGGGGTACCTCGGAAGGCCCTGGATATGTCGGGACGCTATTCATCT 660

QY 742 CCACTCTTCGTGTACAAACAGCGGTATGCCTGTGCGCGGGGAGAACATTAATGCTTCC 801
Db 661 CCACTCTTCGTGTACAAACAGCGGTATGCCTGTGCGCGGGGAGAACATTAATGCTTCC 720

```

```

QY 802 CGAGGAGAGCTGCTGTTCGGCGCAATGATGAGGGTCCCGACGACGACTTCGGC 861
Db 721 CGAGGAGAGCTGCTGTTCGGCGCAATGATGAGGGTCCCGACGACGACTTCGGC 780

QY 862 CTGCTCACCCAGTGCACACACACCATCATGACTATTGGACCTTTGGGATTGGCTGC 921
Db 781 CTGCTCACCCAGTGCACACACACCATCATGACTATTGGACCTTTGGGATTGGCTGC 840

QY 922 TACCTGGCAGTGGTGATACCATCTACTTAGCAACTACACCTTCCGGATTCTCCGTC 981
Db 841 TACCTGGCAGTGGTGATACCATCTACTTAGCAACTACACCTTCCGGATTCTCCGTC 900

QY 982 CTCGAAGTCTTTAAGCCAGAGAGCGCTTCCTACCCGATGGGTGGGATCCCTCCGAT 1041
Db 901 CTCGAAGTCTTTAAGCCAGAGAGCGCTTCCTACCCGATGGGTGGGATCCCTCCGAT 960

QY 1042 CTGTCCCGCACTCCTTAGGCAATTAACACCGCTTCTCTCGTCCCATTCACACCTCAAG 1101
Db 961 CTGTCCCGCACTCCTTAGGCAATTAACACCGCTTCTCTCGTCCCATTCACACCTCAAG 1020

QY 1102 GCAGAAGAGTCACTTGTACGTCCGAGAGAGCGCTTCTGTATGGAA 1149
Db 1021 GCAGAAGAGTCACTTGTACGTCCGAGAGAGCGCTTCTGTATGGAA 1068

RESULT 4
RAD27208
ID AAD27208 standard; DNA; 1068 BP.
XX
AC AAD27208;
XX
DT 09-APR-2002 (first entry)
XX
DE Rat hepatoma H35 cell alpha1-2FucT catalytic domain encoding DNA.
XX
KW Rat; alpha1-2fucosyltransferase; alpha1-2FucT; antisense therapy;
KW galactose beta1-3N-acetylgalactosamine; Galbeta1-3GalNAc; glycolipid;
KW glycoprotein; glycolipoprotein; oligosaccharide; fucosyl-GM1; cancer;
KW gene therapy; oncogenic transformation; cytostatic; ganglioside; GMI;
KW cell transformation; catalytic domain; ds.
XX
OS Rattus norvegicus.
XX
FH Key Location/Qualifiers
CDS 1..1062
FT /*tag= a
FT /product= "Rat hepatoma H35 cell alpha1-2FucT catalytic
FT domain"
FT /note= "CDS does not include start codon"
FT misc_feature
FT 581..1068
FT /*tag= b
FT /note= "Region which overlaps rat FTB"
XX
XX US6329170-B1.
XX
XX 11-DEC-2001.
XX
XX 23-APR-1999; 99US-0298886.
XX
XX 23-APR-1999; 99US-0298886.
XX (NWHO-) NORTHWEST HOSPITAL.
XX
XX Holmes EH, Sherwood AL;
XX
XX WPI; 2002-121132/16.
XX DR P-PSDB; AAE16623.
XX
XX Rat hepatoma H35 cell alpha1-2fucosyltransferase, useful for producing
XX GMI-specific alpha1-2fucosyltransferase enzyme by recombinant
XX techniques and for detecting oncogenic transformation of test tissues -
XX

```



Claim 2; Fig 3; 41pp; English.

The invention relates to rat GM1-specific alpha1-2fucosyltransferase (alpha1-2FucT) enzyme and its corresponding nucleic acid. This nucleic acid is specific for a carbohydrate moiety found in ganglioside GM1, a terminal galactose beta1-3N-acetylgalactosamine (Galbeta1-3GalNAc) saccharide. Alpha1-2FucT DNA is useful for producing rat alpha1-2FucT protein by recombinant techniques. Alpha1-2FucT DNA is useful for the preparative synthesis of fucosyl containing glycolipids, glycoproteins, glycolipoproteins and oligosaccharide, and for preparing fucosyl-GM1. Alpha1-2FucT DNA is useful for detecting oncogenic transformation which involves assaying for changes in expression of alpha1-2 FucT. Since alpha1-2FucT is activated in cell transformation, antisense sequences derived from alpha1-2FucT DNA are useful for inhibiting, suppressing or treating cancer. Alpha1-2FucT DNA is useful in gene therapy and antisense therapy. The present sequence is rat hepatoma H35 cell alpha1-2FucT catalytic domain encoding DNA.

Sequence 1068 BP; 237 A; 308 C; 300 G; 223 T; 0 other;

Query Match 93.0%; Score 1068; DB 24; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 7.3e-288;  
Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

82 CTCACGACGCAATAGTGAAGCTCCAAACCCCTGTGTCAGAGAAAGTAATACCGAIGACGACT 141  
1 CTCACGACGCAATAGTGAAGCTCCAAACCCCTGTGTCAGAGAAAGTAATACCGAIGACGACT 60  
142 CAAATGTCTCTCGGGAACACAGAAAGCCAGAGATGCCAGCGGACGAGCAGCAGATGGG 201  
61 CAAATGTCTCTCGGGAACACAGAAAGCCAGAGATGCCAGCGGACGAGCAGCAGATGGG 120  
202 AATGAGAGCTCCGGGCACTGTTCACGATCAATTCATTCGCGGCTGGGGAAACAGATG 261  
121 AATGAGAGCTCCGGGCACTGTTCACGATCAATTCATTCGCGGCTGGGGAAACAGATG 180  
262 GCGCAATAGCCACACTCTTTCACCTGGCAGGATGAACGAGCGGCTGCGGTTCATCCCC 321  
181 GCGCAATAGCCACACTCTTTCACCTGGCAGGATGAACGAGCGGCTGCGGTTCATCCCC 240  
322 GCATCCATGCACAGCTCTAGCGCCCTCTTCAGGATCAGCTCCCGGTGTACACAGC 381  
241 GCATCCATGCACAGCTCTAGCGCCCTCTTCAGGATCAGCTCCCGGTGTACACAGC 300  
382 GACAGGCCAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGAGCGGT 441  
301 GACAGGCCAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGAGCGGT 360  
442 TACCGCCACATCCGGGACACTTTGTGCGCTTCACGGATACCCCGTCTCTGGACCTTC 501  
361 TACCGCCACATCCGGGACACTTTGTGCGCTTCACGGATACCCCGTCTCTGGACCTTC 420  
502 TACCACCACTCCGCCAGAGATCTGTAAGGAGTTTCACCCCTGCATGACACGCTGGGGAG 561  
421 TACCACCACTCCGCCAGAGATCTGTAAGGAGTTTCACCCCTGCATGACACGCTGGGGAG 480  
562 GAGGCCCAAGCGCTTCTCGCTGGTCTGGCGGTGAATGGAGCCACCGAGTACTTTTGTG 621  
481 GAGGCCCAAGCGCTTCTCGCTGGTCTGGCGGTGAATGGAGCCACCGAGTACTTTTGTG 540  
622 GGTGTCCATGTGCGCCGAGGAGACTATGTGATGTCATGCTTAATGTGGAAGGCGGTG 681  
541 GGTGTCCATGTGCGCCGAGGAGACTATGTGATGTCATGCTTAATGTGGAAGGCGGTG 600  
682 GTGGCTGACCGGGTTACTTGGAAAGCCCTCGGATATGTTCCGGSCACGCTATTCATCT 741  
601 GTGGCTGACCGGGTTACTTGGAAAGCCCTCGGATATGTTCCGGSCACGCTATTCATCT 660  
742 CCAGTCTTCGTGGTTTACAGCAACGCTATGGCCCTGGTCCCGGAGAACATTAATGCTTCC 801  
661 CCAGTCTTCGTGGTTTACAGCAACGCTATGGCCCTGGTCCCGGAGAACATTAATGCTTCC 720  
802 CGAGGAGACGTGGTGTTCGCGGGGCAATGGTATGAGGGGTGCGCCAGCCAGACTTCGCG 861

Db 721 CGAGGAGAGCTGTGTCTCGCGGCATGTATTGAGGGTTCGCCACGACGAGACTTCGCG 780  
Qy 862 CTGCTACCCAGTGGCAACACACACCATCATCATATTGGGACCTTTGGGATTTGGGCTGCC 921  
Db 781 CTGCTACCCAGTGGCAACACACACCATCATCATATTGGGACCTTTGGGATTTGGGCTGCC 840  
Qy 922 TACCTGGCAGGTGGTATACCATCTACTTACCACTACACCATACACCTTCGCGATTCGCTTC 981  
Db 841 TACCTGGCAGGTGGTATACCATCTACTTACCACTACACCATACACCTTCGCGATTCGCTTC 900  
Qy 982 CTCAAAGTCTTTAAGCCAGAGCAGCCCTTCTTACCCGAAATGGGTGGGCATCCCTTCCCGAT 1041  
Db 901 CTCAAAGTCTTTAAGCCAGAGCAGCCCTTCTTACCCGAAATGGGTGGGCATCCCTTCCCGAT 960  
Qy 1042 CTGTCCCACTCTTAAGGCATTAACACACCATTAACACCGGCTGTCTCGGTCCGCTCAAG 1101  
Db 961 CTGTCCCACTCTTAAGGCATTAACACACCGGCTGTCTCGGTCCGCTCAAG 1020  
Qy 1102 GCAAAGGAGTCACTTGTTCAGTTCGCGAGGAGCCCTTCTGATGGAA 1149  
Db 1021 GCAAAGGAGTCACTTGTTCAGTTCGCGAGGAGCCCTTCTGATGGAA 1068  
RESULT 5  
AAV29003  
ID AAV29003 standard; DNA; 1043 BP.  
XX AAV29003;  
AC AAV29003;  
DT 28-AUG-1998 (first entry)  
XX Porcine secretor transferase (FUT2) gene.  
DE Secretor; glycosyltransferase; FUT2; pig; epitope; antigen;  
KW transgenic animal; xenotransplantation; organ transplant; ss.  
XX Sus scrofa.  
XX  
XX Key Location/Qualifiers  
FT CDS 9..1031  
FT /\*tag= a  
XX WO9807837-A1.  
XX  
XX PD 26-FEB-1998.  
XX  
XX PF 22-AUG-1997; 97WO-AU00540.  
XX  
XX PR 23-AUG-1996; 96AU-0001823.  
XX  
XX (AUST-) AUSTIN RES INST.  
XX  
XX PI McKenzie IFC, Sandrin MS;  
XX  
XX DR WPI; 1998-169148/15.  
XX P-PSDB; AAW37855.  
XX  
XX PT Nucleic acid encoding glycosyltransferase able to compete with  
PT second such enzyme - particularly used to reduce expression of  
PT unwanted carbohydrate epitope(s) on tissues intended for  
PT transplantation  
XX  
XX PS Claim 6; Fig 1A-B; 40pp; English.  
XX  
XX This nucleotide sequence, designated clone pSe16.1, codes for  
CC porcine secretor (Se) glycosyltransferase (see AAW37855), an enzyme  
CC that has high affinity for type I and type III substrates. It was  
CC isolated from a pig liver genomic library using full-length human  
CC Se2 (FUT2) cDNA as probe. The nucleotide sequence shows about 36%  
CC homology with human FUT1. Introduction of glycosyltransferase  
CC nucleic acid, such as porcine Se nucleic acid, into a tissue  
CC results in reduced expression of unwanted carbohydrate epitopes on

the tissue, especially porcine heart, liver, kidney or pancreas, rendering it more suitable for transplantation, i.e. less immunogenic and of increased immunological acceptability. A claimed method of producing a cell from a donor species that is immunologically acceptable to a recipient species involves reducing levels of carbohydrate on the donor cell that causes it to be recognised as non-self by the recipient by causing a nucleic acid for a glycosyltransferase such as Se to be expressed in the cell. Expression units, such as retroviral packaging or producer cells, containing Se nucleic acids can be used in gene therapy.

SQ Sequence 1043 BP; 189 A; 352 C; 306 G; 196 T; 0 other;

Query Match 54.8%; Score 530.2; DB 19; Length 1043;

Best Local Similarity 76.9%; Pred. No. 1.1e-165;

Matches 814; Conservative 0; Mismatches 203; Indels 42; Gaps 2;

1	ATGGCCAGCGCCAGGTTCCCTTTCTCTCTTCCCTCTGCCCACTTCCTCCTCATCTTTGTGCTTC	60	
QY			
9	ATGTCACATGCATGAGGATCCTTTCTTCTCCCAAGGTCCTCTCATCTCTTTGTGCTTC	68	
Db			
61	GTGACTTCCACCATCTCCACTCCACAGCGAATGTGAAGCTCAACACCCCTGTCCAGAC	120	
QY			
69	ACGGATTCACCATATTTCACCTTCACAGAGATGTTGAAGATTCACCCACGT	123	
Db			
121	AAGGAATTACCGATGACGATTCAMATTCCTTCGGGAACACAGAAAGCCACAGATACGGA	180	
QY			
124	-GGAGATTACAGATGGTCAACAGAGTCAACACAGAGAGCCCTCGAGCC-	172	
Db			
181	CGGACACGAGCAGCATGGGAATGGAGAGCTCGGGGCATGTCAGATCAATTCATT	240	
QY			
173	-----CCAGTGAAGGCATGTGGACGATCAATGCCATC	206	
Db			
241	GGCGGCTGGGAACACAGATGGGGAATAGCCACACTCTTTCGACTGGCCAGGATGAAC	300	
QY			
207	GGCGGCTGGGAACACAGATGGGGAGTACGCACCTGTACGGCTGGCCAGAGTAC	266	
Db			
301	GGAGGGCTTGGTTTCATCCCGGCATCCATGCACAACGCTCTACGGCCCATPTTCAGATC	360	
QY			
287	GGCGGGCGGGCTTCATCCGGCCGAGATGCACAGACGCTGGCCGCCCATCTTCAGGATC	326	
Db			
361	AGCCTCCGGTGTTCACACAGGCACACGGCCAAAAGATCCATGGCAGAGATACCATCTC	420	
QY			
327	ACCTCCGGTCTCTGCAGCCAGCACGCGCCGAGGATCCCTGGCAGAACTACCACTG	386	
Db			
421	AACGACTGGATGGAGGAGCGTTACCGGCCATTCGCGGACACTTTGTGGCTTTCAGGGA	480	
QY			
387	AACGACTGGATGGAGGAGCGGTACCGCCACATCCGGGGAGTACGTGGCCCTCACGGC	446	
Db			
481	TACCGTGCTCCTCGACCTTCTACCAACACTCGGCCACAGATCCTGAAGAGTTCAAC	540	
QY			
447	TACCCCTGCTCCTGGACCTTCTACCAACCTCGCACCGAGATCCTCGGGAGTTCAAC	506	
Db			
541	CTGCATGACCACGTGCGGGAGAGGCCACGGCTTCTCGTGCTCTGGGGTGAATGG	600	
QY			
507	CTGCATACCAACGTTGCGGAGGAGGCCAGGATTCCTCGGGGTCTGGGTGAACGG	566	
Db			
601	ACCAAGCCGAGTACTTTGTGGGHTCCATGTGGCGGAGGGGACTATGTGATGTCAATG	660	
QY			
567	AGCCGACCGATACCTACGTGGGGGTGCACGTGGCCGGGGGACTACGTGACGCGTATG	626	
Db			
661	CCTAATGTGTGAAGCGGTGGGCTGACCGGGGTTACCTGGAAGGCCCTGGATATG	720	
QY			
627	CCCAAGTGTGGAAGGCGTGTGTGGCCGACCGCGGTACCTGGAGCAGGCCCTGGACTGG	686	
Db			
721	TTCCGGGCACCATTCATCTCCAGTCTTCGTGTGTACAGCAACGGTATGGCTGGTGC	780	
QY			
687	TTCCGGGCTCGCTACCGCTCCCGCTTTTGTGTCTCCAGCAACGGCATGGCTGGTGT	746	
Db			
781	CGGAGAACATTAATGTTCCCGAGGAGACGTGGTGTTCGCGGCCAATGATTTGAGGG	840	
QY			
747	CGGGAACATCAATGCTCGCGGGCGGATGTGTGTTCGCGGCAATGGATTCAGGCG	806	
Db			



















P-PSDB; AAY79302.  
Improving weight gain in swine using swine genetically resistant  
Escherichia coli and feeding swine high levels of plant based protein  
Disclosure; Fig 1; 33pp; English.  
This is the nucleotide sequence of the pig FUT1 gene encoding  
alpha-1-2 fucosyltransferase (see AAY79302). A polymorphism at  
position 307 of the coding region is associated with susceptibility  
to F18 Escherichia coli colonization; pigs homozygous for adenine  
at position 307 are resistant to colonization, while heterozygous  
animals and animals homozygous for guanine at position 307 are  
susceptible to colonization. A claimed method for improving weight  
gain in pigs involves selecting animals that are genetically  
resistant to E. coli colonization and feeding these animals high  
levels of plant-based protein concentrate. A claimed method for  
preventing F18 E. coli colonization in swine, especially swine that  
are genetically susceptible to F18 E. coli colonization, involves  
replacing some or all of the plant-based proteins in the diet with  
animal-based proteins. The polymorphism in the FUT1 gene is also  
useful for developing drugs to treat swine that have E. coli  
associated disease. The polymorphism can be detected using  
PCR-RFLP tests (see also AAZ94418-19).  
Sequence 1269 BP; 219 A; 413 C; 362 G; 275 T; 0 other;  
Query Match 36.9%; Score 423.8; DB 21; Length 1269;  
Best Local Similarity 69.6%; Pred. No. 5e-108;  
Matches 591; Conservative 0; Mismatches 252; Indels 6; Gaps 1;  
217 GGCATGTTACAGTCAATTCATCCATGGCGGTGGGGAACACAGATGGGGAATACGCCACA 276  
243 GGGACCTGGACTATTACCGGATGGCGGTTGGGAACACAGATGGGACATATGCCACG 302  
277 CTCCTTGGACCTGGGAGATGAACGAGCGGTGGTTCATCCCGCATCCATGACACAAAC 336  
303 CTGCTGGCCCTGGGAGCTCAACGGCGCCAGGCGCTTCATCCAGCCTGCCATGACGCC 362  
337 GCTCTAGCGCCCACTCAGATCAGCTCCCGGTGGTGTACACGGGACACGCCAAAG 396  
363 GTCCTGGCCCGGTGTCGGATCAGCTGCTGCTGGCGCGGAGGTAGACAGGCAC 422  
397 ATCCATGGCAGAAATTACATCTCAACACTGGATGGAGGAGCGTTACCGCAACATTCGG 456  
423 GCTCTGGCGGGAGCTGGAGCTCAACAGCTGGATGTCGGAGGATTATGCCACATTAAAG 482  
457 GGACACTTGTGGGTTACGGGATACCGGTGCTGCTGGACCTTCTACACACACCTGGCG 516  
483 GAGCCCTGGCTGAAGCTACACGGGCTTCCCTGCTGGCGCGGAGGTAGACAGGCAC 542  
517 CCAGAGATCCTGAAGGAGTTCACCTGCATGACACAGCTGGCGGAGGAGGCCAGGCCTTC 576  
543 GAGCAGATCCGACGAGTTCACCTGCACGACACACCTTCGGCAGAGGCCCGGAGGGTA 602  
577 CTGCCT-----GGTCTCGGGTGAATGGGAGCGAGCGAGTACTTTGTGGGTGTCAT 630  
603 CTGAGTCAAGTCCGCTACACCCGCACAGGGGACGCCGCCACCTTCGTGGGGTCCAC 662  
631 GTGGCGGAGGGGACTATGTCATGTCATGCTTAATGTGGAGGGCGTGGTGGCTGAC 690  
663 GTGGCGCGGGGACTATCTCGGTGATGTCGCAAGCGCTGGAAGGGGGTGGTGGTAC 722  
691 CGGGTTHACCTGGAAAGGCCCTGATATGTCGGGCAACGCTATTCATCTCCAGCTTC 750  
723 GGCCTTACCTCCAGAGGCTATGACGTGTTCCGGGCCCATAGAGCCCGGCTTT 782  
751 GTGGTACAAGCAACGGTATGCTGGTGGCGGGAGAACATTAATGCTTCCCGAGGAC 810  
783 GTGGTCACACACGAGCGATGGATGGTGGCGGGAGAACATCGACACCTCCCGGGGAC 842  
811 GTGGTGTTCGGGGCAATGTTATGAGGGGTGCGCAGCAAGGACTTCGCGCTGCTCACC 870

Db 843 GTGATCTTGTGGGATGGCGGNGCGCGCCGAGGACATTGGCTGCTGGTS 902  
Qy 871 CAGTCAACACACCATCATGACTATTGGGACTTTGGGACTTTGGGCTCCCTACCTGGCA 930  
Db 903 CAGTCAACACACCATCATGACTATTGGGACTTTGGGACTTTGGGCTCCCTACCTGGCT 962  
Qy 931 GTGTGTATACCATCTACTTAGCCACTACACCTTCCGGATTCCTCCGTCCCAAGTC 990  
Db 963 GTGGAGATACCATCTACTTAGCTAACTTCACTTCCGCTCCGCTCCGCTTCTGAAGATC 1022  
Qy 991 TTAAAGCCAGAGGAGCGCTTCTTACCCGAATGGTGGGATCCCTGCCGATCTGTCGCCA 1050  
Db 1023 TTAAACCGAGGCGCTTCTTCTGCCGAGTGGTGGCATTAATGACACTTGTCTCCA 1082  
Qy 1051 CTCCTTAAG 1059  
Db 1083 CTCAGATG 1091  
RESULT 15  
AAI72831  
ID AAI72831 standard; cDNA; 1269 BP.  
XX  
AC AAI72831;  
DT 22-JUL-2002 (first entry)  
XX  
DE FUT1 cDNA.  
KW Gen; pig; swine; alpha (1,2) fucosyltransferase; FUT1; weight gain;  
F18; Escherichia coli; ss.  
XX  
OS Sus scrofa.  
FH Key Location/Qualifiers  
FT CDS 9..1106  
FT /tag= a  
FT /product= "FUT1"  
FT mutation 315  
FT /tag= b  
FT /phenotype= "Confers resistance to F18 E. coli"  
XX US6355859-B1.  
PN  
PD 12-MAR-2002.  
XX  
PF 18-SEP-1998; 980S-0151592.  
XX  
PR 20-MAY-1997; 970S-047181P.  
XX  
PA (BIOT-) BIOTECHNOLOGY RES & DEV CORP.  
PA (USDA ) US SEC OF AGRIC.  
XX  
PI Bosworth B, Ridpath J, Wiseman B;  
XX  
DR WPI; 2002-391652/42.  
XX  
XX P-PSDB; AAB47995.  
XX  
XX Feeding swine genetically susceptible to F18 Escherichia coli  
XX colonization with a low plant protein based diet increases weight gain  
XX and lowers intestinal disease associated with E. coli infection  
PS Claim 1; Column 13-18; 9pp; English.  
XX  
XX This sequence represents the swine alpha (1,2) fucosyltransferase  
XX (FUT1) gene. A FUT1 gene in which there is a base other than adenine  
XX at position 307, may be used for improving weight gain in swine that  
XX are genetically susceptible to F18 Escherichia coli. The weight gain  
XX may be activated by feeding a diet of at least 40% animal based  
XX proteins. The feeding method is used to control F18 E. coli associated  
XX intestinal disease in swine.

SQ Sequence 1269 BP; 219 A; 414 C; 361 G; 275 T; 0 other;									
Query Match 36.9%; Score 423.8; DB 24; Length 1269;									
Best Local Similarity 69.6%; Pred. No. 5e-108;									
Matches 591; Conservative 0; Mismatches 252; Indels 6; Gaps 1;									
Qy	217	GGGATGTTACAGATCAATTCATTCGCGGCTGGGAACAGATGGCGAATACGCCACA	276						
Db	243	GGGACCTGGGACTATTACCGGATGGCGGTTTGGGAACAGATGGACAGTATGCCACG	302						
Qy	277	CTCTTTCATCGCCAGGATGACGAGGCTTGGCTTATCCCGCATCCATGCACAAC	336						
Db	303	CTGCTGGCCCTGGCGACGCTCAACGGCGCCAGGCTTCATCCAGCTGCATGCACGCC	362						
Qy	337	GCCTAGCGCCCTCTTCAGGATCAGCTCCCGGTGTTACACAGCGACACGGCCAAAAG	396						
Db	363	GTCTGSCCCCGCTGTTCGCGATCACCTGCTGCTCTGGGCCGAGGTAGACAGGCAC	422						
Qy	397	ATCCATGGGAGAAATACCATCTCAAGACTGGATGGAGAGCCTTACCGCCACATTCCG	456						
Db	423	GCTCCTTGGCGGAGCTGGAGCTTCAGACTGGATGTCGAGGATTATGCCACTTAAAG	482						
Qy	457	GGACACTTTGTGCGCTTCACGGATACCGGTGCTCTCGGACCTTCTACACACCTGGC	516						
Db	483	GAGCCCTGGGTGAAGCTCACCGGCTTCCCTGCTCTCTGACCTTCTTCCACCACTCCG	542						
Qy	517	CCAGAGATCCTGAAGAGTTCACTCATGACCACTGCGGAGGAGGCCACGGCCTTC	576						
Db	543	GAGCAGATCCGACGAGTTCACTCATGACCACTGCGGAGGAGGCCACGGCCTTC	602						
Qy	577	CTGGCT-----GGTCTGGGGTGAATGGAGGAGCGGAGTACTTTTGGGTGTCCAT	630						
Db	603	CTGAGTCAGTTCCGCTTACCCCGCACAGGAGGAGCGCCCGCCAGCACTTCGTGGGGTCCAC	662						
Qy	631	GTGCGCGGAGGGACTATGTCATGTCATGCTTAATGTGGRAGGGCGTGGTGGCTGAC	690						
Db	663	GTGCGCGCGGGGACTATGTCGCTGTATGCCCAAGCGCTGGAAGGGGGTGGTGGGTGAC	722						
Qy	691	CGGGTTACCTGGAAGAGGCGCTGGATATGTCGGGCGACGCTATTCTCCAGTCTTC	750						
Db	723	GGCCCTTACCTCCAGCAGGCTATGAGCTGTTCCGGGCCGATACGAGCCCGCTCTTT	782						
Qy	751	GTGTTTACAGCAAGGTATGGCTGTGTCGGGAGAACATTAATGTTCCCGAGGAGAC	810						
Db	783	GTGTTACCAAGCAAGGATGGAGTGTGTCGGGAGAACATCGACACCTCCCGGGGGAC	842						
Qy	811	GTGCTGTTCGGGCAATGTTATGAGGGGTGCGCCAGCCAGGACTTCGGCTGCTCACC	870						
Db	843	GTGATCTTTGCTGGCGATGGCGGGAGGCGCGCCCGCCAGGACTTTGCGCTGCTGTG	902						
Qy	871	CAGTCAACCAACACATCATGACTATTGGGACCTTTGGGATTGGGCTGCGCTACCTGGCA	930						
Db	903	CAGTCAACCAACACATCATGACTATTGGGACCTTTGGGATTGGGCTGCGCTACCTGGCT	962						
Qy	931	GGTGGTATACATCTACTTAGCAACTACACCTTCCGGATTTCCCGTTCTCTCAAAGTC	990						
Db	963	GGTGGATACCATCTACTTGGTAACTTCACTCCCTGCGCACTTCCAGCTTCTGAAGATC	1022						
Qy	991	TTTAGCCAGAGGAGCCTTCTTACCCCAATGGTGGGCATCCCTGCGGATCTGTCCCA	1050						
Db	1023	TTTAAACCGAGGTGCTTCTCTGCGGAGTGGGTGGGCAATTAATGCAGACTTGTCTCCA	1082						
Qy	1051	CTCCTTAAG	1059						
Db	1083	CTCCAGATG	1091						

Search completed: May 26, 2003, 12:08:24  
Job time : 302.004 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 05:50:26 ; Search time 62.7104 seconds  
(without alignments)  
5819.028 Million cell updates/sec

Title: US-10-040-863-7

Perfect score: 1149

Sequence: 1 atggccagcgcccggttcc.....gaagagcctctgatggaaa 1149

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCrUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1149	100.0	1149	4	US-09-298-886-7
2	1068	93.0	1068	4	US-09-298-886-9
3	630.2	54.8	1043	4	US-09-254-077A-5
4	613.8	53.4	2115	1	US-08-395-800A-7
5	516.6	45.0	1144	1	US-08-395-800A-1
6	423.8	36.9	1269	4	US-09-151-592-1
7	408.2	35.5	1155	1	US-08-208-889A-1
8	408.2	35.5	1155	2	US-08-433-271-1
9	405	35.2	999	5	PCR-US91-00899-10
10	405	35.2	1136	1	US-08-395-800A-9
11	405	35.2	1174	5	PCR-US95-07554-3
12	405	35.2	1199	1	US-08-395-800A-5
13	405	35.2	2268	4	US-08-675-773B-4
14	405	35.2	3373	1	US-08-273-411-2
15	405	35.2	3791	4	US-08-675-773B-3
16	405	35.2	8174	1	US-07-914-281-5
17	405	35.2	8174	1	US-08-393-246-5
18	405	35.2	8174	1	US-08-526-058A-5
19	405	35.2	8174	2	US-08-696-731-5
20	405	35.2	8174	4	US-09-042-531-5
21	405	35.2	8174	5	PCR-US91-00899-3
22	403.4	35.1	1155	1	US-08-434-151-1
23	403.4	35.1	1155	2	US-08-715-259-1
24	72.8	6.3	100	4	US-09-298-886-29
25	71.2	6.2	100	4	US-09-298-886-23
26	71.2	6.2	100	4	US-09-298-886-26
27	69.6	6.1	100	4	US-09-298-886-20

```

28 63.2 5.5 100 4 US-09-298-886-14 Sequence 14, Appl
29 63.2 5.5 100 4 US-09-298-886-17 Sequence 17, Appl
30 55.8 4.9 100 4 US-09-298-886-15 Sequence 15, Appl
31 49.8 4.3 79 4 US-09-298-886-27 Sequence 27, Appl
32 38.4 3.3 50 4 US-09-298-886-16 Sequence 16, Appl
33 37.6 3.3 1569 2 US-08-997-080-113 Sequence 113, Appl
34 37.6 3.3 1569 2 US-08-997-362-113 Sequence 113, Appl
35 37.6 3.3 1569 4 US-09-095-855-113 Sequence 113, Appl
36 37.6 3.3 1569 4 US-09-324-542-113 Sequence 113, Appl
37 37.6 3.3 1569 4 US-09-203-426-113 Sequence 113, Appl
38 37.6 3.3 1626 2 US-08-997-080-159 Sequence 159, Appl
39 37.6 3.3 1626 2 US-08-997-362-159 Sequence 159, Appl
40 37.6 3.3 1626 4 US-09-095-855-159 Sequence 159, Appl
41 37.6 3.3 1626 4 US-09-324-542-159 Sequence 159, Appl
42 37.6 3.3 1626 4 US-09-203-426-159 Sequence 159, Appl
43 37.4 3.3 50 4 US-09-298-886-28 Sequence 28, Appl
44 37 3.2 50 4 US-09-298-886-25 Sequence 25, Appl
45 36.8 3.2 50 4 US-09-298-886-19 Sequence 19, Appl

```

#### ALIGNMENTS

RESULT 1

US-09-298-886-7

; Sequence 7, Application US/09298886

; Patent No. 6329170

; GENERAL INFORMATION:

; APPLICANT: ERIC H. Holmes et al.

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE

; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 8511-029

; CURRENT APPLICATION NUMBER: US/09/298,886

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 7

; LENGTH: 1149

; TYPE: DNA

; ORGANISM: Rattus norvegicus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1143)

US-09-298-886-7

```

Query Match      100.0%; Score 1149; DB 4; Length 1149;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCCAGCGCCAGGTTCCCTTCTCCTTTCCTTCCTTCGGCCACTTCCTCATCTTTGCTTC 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 ATGGCCAGCGCCAGGTTCCCTTCTCCTTTCCTTCCTTCGGCCACTTCCTCATCTTTGCTTC 60

Qy 61 GTGACTTCCACCATCATCCACCTCCAGCAGCGAATAGTGAAGCTCCACCCCTGTCAGAG 120
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 GTGACTTCCACCATCATCCACCTCCAGCAGCGAATAGTGAAGCTCCACCCCTGTCAGAG 120

Qy 121 AAGGAATACCGATGACGACTCAATGTCTCGGGAACACAGAACGCCAGAGATGGA 180
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 AAGGAATACCGATGACGACTCAATGTCTCGGGAACACAGAACGCCAGAGATGGA 180

Qy 181 CGGGACAGCAGCAGCATGGGAATGGAGCTGGGGGCAATTCACGATCAATTCATT 240
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CGGGACAGCAGCAGCATGGGAATGGAGCTGGGGGCAATTCACGATCAATTCATT 240

Qy 241 GCGCGGCTGGGGAACAGATGGCGCAATACGCCACACTTTTGCACCTGCCAGGATGAAC 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 GCGCGGCTGGGGAACAGATGGCGCAATACGCCACACTTTTGCACCTGCCAGGATGAAC 300

Qy 301 GGACGGCTGGCTTCATCCCGCATCCATGACACACGCTTAGCGCCCATCTTTCAGGATC 360
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 GGACGGCTGGCTTCATCCCGCATCCATGACACACGCTTAGCGCCCATCTTTCAGGATC 360

```









```

; SEQ ID NO 1
;
; LENGTH: 1269
;
; TYPE: DNA
;
; ORGANISM: Swine
;
; FEATURE:
;
; NAME/KEY: CDS
;
; LOCATION: (9)..(1103)
;
; FEATURE:
;
; OTHER INFORMATION: FUTI
US-09-151-592-1

```

Query Match 35.9%; Score 423.8; DB 4; Length 1269;  
Best Local Similarity 59.6%; Pred. No. 7.9e-114;  
Matches 591; Conservative 0; Mismatches 252; Indels 6; Gaps 1;

QY	217	GGCATGTTACAGGATCAATTCATTTGGCCGGCTGGGAAACAGATGGGGGAATACGGCACA	276
Db	243	GGACCTGGACTATTACCGGGATGGCCGTTTGGNACAGATGGGACAGTATGCCACG	302
QY	277	CTCTTTGCACTGSCCAGGATGAACGGACGGCTTGCCTTCATCCCGCATCCATGCAACAAC	336
Db	303	CTGTGGCCCTGGCGAGCTCAACGGCGCCAGGCCCTTCATCCAGGCTGCATATGCACGCC	362
QY	337	GCCTAGCGCCCATCTCAGGATCAGCCTCCCGGTGTACACAGGACACGGGCCCAAAAG	396
Db	363	GTCTGGCCCCGTGTTCCGCATCAAGCTGCCCTGCTCTGGCGCCGAGTAGACAGGCAC	422
QY	397	ATCCCATGGCAGAAATACCATCTCAAGACTGGATGGAGGAGGTTACGCCACATTCG	456
Db	423	GTCTCTGGCGGGAGCTGGAGCTTCAGGACTGGATGTCGAGGATTAATGCCACTTAAAG	482
QY	457	GGACACTTTGTGGCTTCACGGGAAACCCGTGCTCTGGACCTTTCACACACACTTGGCC	516
Db	483	GAGCCCTGGGTGAAGCTCACCGGCTCCCTGCTCTGGACCTTCTTCCACCACTCCGG	542
QY	517	CCAGAGATCTGAAGGAGTTACCCCTGCATGACACAGTGGGGAGGAGGCCAGGCGCTTC	576
Db	543	GACGAGATCCGACGAGTTTACCTCTGCAGCACACCTTCGGCAAGAGGCCACAGGGGTA	602
QY	577	CTCGCT-----GGTCTGGCGGGTGAATGGAGGCAGCCGAGTACTTTGTGGTGCCAT	630
Db	603	CTGAGTCAGTCCGTCTACCCGACACAGGGACCGGCCAGCACTTCGTGGGGTCCAC	662
QY	631	GTGGCGGAGGGACATATGTGCATGTCATGCCCTAATGTGTGGAGAGCGCTGGTGGCTGAC	690
Db	663	GTGGCGCGGGGACTATCTCGCTGTATGCCCAAGCGCTGGAGGGGCTGGTGGGTGAC	722
QY	691	CGGGGTTACCTGAAAGAGCCCTGGATATGTTCCGGGCAGGCTATTATCTCCAGTCTTC	750
Db	723	GGCGCTTACCTCCAGCAGGCTATGGACTGGTTCGGGCGCGATACGAAAGCCCGCTTT	782
QY	751	GTGGTTACAGCAACGGTATGGCGTGTGGCCGGGAGAACATTAATGTTCCGAGAGAC	810
Db	783	GTGGTCACCAAGCAACGGCATGGAGTGTGGCGGAAGAACATCGACACCTCCCGGGGGAC	842
QY	811	GTGGTGTCGCGGCAATGATATGAGGGGTGCCAGCAAGGACATTCGGGCTGCTCAAC	870
Db	843	GTGATCTTTGCTGGCGATGGCGGAGGCGCGCCCGCAGGAGCTTTGGCTGCTGGTG	902
QY	871	CAGTGAACACACCACTCATGACTATTGGACACTTTGGATTTGGGCTGCTACCTGGCA	930
Db	903	CAGTGAACACACCACTCATGACCATTTGGCACTTCGGCTCTGGCGCGGCTACCTGGCT	962
QY	931	GGTGGTGATACATCTACTTATGGCACTACACCTTCCGGATTCCTCGTTCCTCAAGTC	990
Db	963	GGTGGGATACCATCTACTTGGTAACTTCAACCTGCCACTTCCAGCTTCCCTGAAGATC	1022
QY	991	TTTAAAGCCAGAGCAGCTTCTCCTACCGAATGGTGGGATCCCTCCGATCTGTCCCCA	1050
Db	1023	TTTAAAGCCAGAGCTGCTTCTCTCCCGAGTGGGAGGCAATTAATGCAGACTTGTCTCCA	1082
QY	1051	CTCCTTAAAG	1059

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.0

Db 1083 CTCAGATG 1091

RESULT 7

US-06-208-889A-1  
Sequence 1, Application US/08208889A  
Patent No. 5750176  
GENERAL INFORMATION:  
APPLICANT: Prieto, P., Smith, D., Cummings, R., Kopchick, J., Mukerji, P.,  
APPLICANT: Moremen, K., Pierce, J.  
TITLE OF INVENTION: Transgenic Production of Oligosaccharides and  
TITLE OF INVENTION: Glycoconjugates  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Donald O. Nickey  
ADDRESSEE: ROSS Products Division  
ADDRESSEE: Abbott Laboratories  
STREET: 625 Cleveland Avenue  
CITY: Columbus  
STATE: Ohio  
COUNTRY: United States of America  
ZIP: 43215  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS Version 6.21  
SOFTWARE: WordPerfect Version 6.0a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/208,889A  
FILING DATE: 09-MAR-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA: No. 5750176 applicable  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (614) 624-7080  
TELEFAX: (614) 624-3074  
TELEX: No. 5750176e  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1155 base pairs  
TYPE: Nucleic acid  
STRANDEDNESS: Single  
TOPOLOGY: Unknown  
MOLECULE TYPE: Cloned cDNA representing the product of a  
MOLECULE TYPE: Human genomic DNA segment  
DESCRIPTION: GDP-L-fucose- -D-galactoside 2-alpha-fucosyl-transferase  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE: Entire amino acid sequence provided.  
ORIGINAL SOURCE: Human Epidermal Carcinoma Cell line  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE: Human Epidermal Carcinoma Cell line  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 19  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD: DNA sequencing and restriction analysis  
OTHER INFORMATION: The encoded product of nucleotide SEQ ID NO:  
OTHER INFORMATION: 1: is the enzyme, GDP-L-fucose- -D-galactoside 2-alpha-fucosyl-  
OTHER INFORMATION: which has the amino acid sequence described in SEQ. ID NO:1:  
OTHER INFORMATION: is responsible for the synthesis of 2'-fucosyllactose.

PUBLICATION INFORMATION:  
AUTHORS: V.P. Rajan, et al.  
TITLE: Journal of Biological Chemistry  
JOURNAL: Journal of Biological Chemistry  
VOLUME: 264  
ISSUE: 11158 - 11167  
PAGES: 1989  
DATE: 1989  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-08-208-889A-1  
Query Match 35.5%; Score 408.2; DB 1; Length 1155;  
Best Local Similarity 68.7%; Pred. No. 2.6e-109;  
Matches 579; Conservative 0; Mismatches 258; Indels 6; Gaps 1;  
Qy 217 GGCATGTTTCAGATCAATTCATTGGCGGGTGGGAACACAGATGGCGAATACGCCACA 276  
Db 272 GGCACCTGGACTGCTACCCCAANTGGCGGTTTGTATCAGATGGACAGTATGCCACG 331  
Qy 277 CTCTTTGCACCTGGCCAGAGTGAACGGAGCGCTTCGTTTCATCCCGCATCCATGCACAAC 336  
Db 332 CTGCTGGCTCTGGCCAGCTCAACGGCGCGCGGCTTTATCTGCTGCGCATGCATGCC 391  
Qy 337 GCTCTAGCGCCCATCTTCAGATCAGCTCCCGGTGTTTACACAGCGACACGGCCAAAAAG 396  
Db 392 GCCCTGGCCCGGTATTCCCGCATCACCTGCCGIGTGGCCCCCAGAGTGGACGGCGC 451  
Qy 397 ATCCATGGCAGAATTACCATCTCAACGACTGGATGGAGAGCGTTACCGCCACATCCG 456  
Db 452 ACGCGTGGGGAGCTGCAGCTTCACGACGTAIGTCGGAGGAGTACGGCGGACTTGAGA 511  
Qy 457 GGACACTTTGTGCGCTTCACGGGATACCGTGCTCCCTGGACCTTCTACCCACCCTGGC 516  
Db 512 GATCCTTTCTGGAAGCTCTGCTGCTCCCTGCTCTTGGACTTTCTTCCACACTCTCGG 571  
Qy 517 CCAGATCTCTGAAGGAGTTTCACTGATGATGATGATGATGATGATGATGATGATGATG 576  
Db 572 GACAGATCCGACAGAGTTCACCTTCACACACACCTTCGGAAGAGAGGCGCAGGTG 631  
Qy 577 CTGCTGGTCTGCGGGTG-----AATGGAGCGAGCGGAGTACTTTTGGTGGTGTCCAT 630  
Db 632 CTGCTGAGTCTCCCTGGCGGCGACAGGGAGCGCGCGGCGACCTTTGTGGCGTCCAC 691  
Qy 631 GTGGCGGAGGGGACTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 690  
Db 692 GTGGCGGTGGGAGTATCTGAGGTTATGCTGAGGTTATGCTGAGGTTATGCTGAGGTT 751  
Qy 691 CGGGTTACCTGSAAGGCGCTGATATGTTCCGGCACGCTATTCATCTCCAGTCTTC 750  
Db 752 AGCGCTACCTCCGCGAGGCGATGAGTGGTTCCGGCACGCGCACGAGCGCGCTTTC 811  
Qy 751 GTGTTACAGCAACGGTATGGCTGTGGCGGAGAGACATTAATGCTTCCGAGGAGAC 810  
Db 812 GTGTTACAGCAACGGTATGGTGTGGTGTAAAGAAACATCGACACTCCCGAGGGCAT 871  
Qy 811 GTGTTTTCGGGCGCAATGTTTGGGTTGCGCGAGGAGTTCGGGCTCTCTCACC 870  
Db 872 GTGAGTTTCTGGCGATGGACAGGAGCTACCGTGGAAAGACTTTGCCCTGCTCCA 931  
Qy 871 CAGTGCAACCAACCATCATGACTATTGGGAGCTTTGGGATTTGGGCTGCTACCTGGCA 930  
Db 932 CAGTGCAACCAACCATCATGACTATTGGGAGCTTTGGGATTTGGGCTGCTACCTGGCA 991  
Qy 931 GGTGGTATACCATCTACTTAGCCAACTACACCTTCCGGATTCCTGCTTCTCAAGTCC 990  
Db 992 GCGGAGACACTGCTACCTGGCCAACTTACCTTGCACAGACTCTGAGTTCTCTGAATC 1051  
Qy 991 TTTAGCCAGGAGGAGCTTCTTACCCGAATGGTGGGATCCCTGCGGATCTGCTCCCA 1050  
Db 1052 TTTAAGCCGAGGAGGAGCTTCTTCCCGGAGTGGGTGGGCAATTAATGAGACTTCTCTCCA 1111



Db 992 GCGGAGACACTGTCTACCTGGCCAACTTCAACCTGCCAGACTGTGAGTTCCTGAAGATC 1051  
Qy 991 TTAAAGCCAGGAGCGCTTCTTACCGGATGGTGGGATCCCTGCCGATCTGTCCCA 1050  
Db 1052 TTAAAGCCAGGAGCGCTTCTTACCGGATGGTGGGATCCCTGCCGATCTGTCCCA 1111  
Qy 1051 CTC 1053  
Db 1112 CTC 1114

RESULT 9  
PCT-US91-00899-10  
; Sequence 10, Application PC/TUS9100899  
; GENERAL INFORMATION:  
; APPLICANT: Lowe, John B.  
; TITLE OF INVENTION: Method and Products For the Synthesis of  
; TITLE OF INVENTION: Oligosaccharide Structures on Glycoproteins, Glycolipids,  
; TITLE OF INVENTION: or as Free Molecules, and For the Isolation of Cloned  
; TITLE OF INVENTION: Genetic Sequences That Determine These Structures  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/00899  
; FILING DATE: 19910214  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye Ph.D., Jean-Paul  
; REGISTRATION NUMBER: 31,451  
; REFERENCE/DOCKET NUMBER: 2363-021-55 PCT  
; TELEPHONE: (703)521-5940  
; TELEFAX: (703)486-2347  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 999 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; TISSUE TYPE: Blood  
PCT-US91-00899-10

Query Match 35.2%; Score 405; DB 5; Length 999;  
Best Local Similarity 68.4%; Pred. No. 2.1e-108;  
Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;

Qy 217 GGCATGTTCCAGTCAATTCATTCATGGCGGTGGGAACACAGATGGGGAATAGCCACA 276  
Db 139 GGCACCTGGATGTCTACCCCAATGGCGGTGTGTAATCAGATGGGACATATGCCACG 198  
Qy 277 CTCCTTTCAGTCGGCAGGATGAACGAGCGGCTTGGTTATCCCGCGCATCCATGCACAA 336  
Db 199 CTGCTGGCTCTGGCCACACTACAGCGCGCGGGCTTTATCCCTGCTGCCATGCATGCC 258  
Qy 337 GCTCTAGCGGCCATGCTTCCAGTACAGCTCCCGGTGTGTACAGGACACGCGCAAAAG 396

Db 259 GCCTTGGCCCGGTATTCCTGCATCACCTGCCCTGCCCTGGCCAGAGTGAGAGCCGC 318  
Qy 397 ATCCATGGCAGATTTACCATCTCAAGACTTGGATGGAGAGCGCTTACCGCCACATTCG 456  
Db 319 AGCCGTGGCGGAGCTGCAGCTTCACGACTTGAATGTCGGAGGAGTACGGGACTTGAGA 378  
Qy 457 GGACACTTTTGGCGCTTCAGGGATACCCGCTGCTTCCCTGGACCTTCTACCAACCACTCCG 516  
Db 379 GATCCTTTCTGAAGCTCTCTGGCTTCCCTGCTTCTTGGACTTCTTCCCACTCCCG 438  
Qy 517 CCAGAGATCTCTGAAGAGTTTCAACCTGCATGACACGTCGGGAGAGAGGCCGCTTC 575  
Db 439 GAACAGATCCGAGAGATTTCAACCTGCACACACCTTCGGGAAGAGGCGACAGTGTG 498  
Qy 577 CTGCTGTCTGGGGTG-----AATGGAGCGAGCGGAGTACTTTTGGGTGTCCAT 630  
Db 499 CTGGGTGAGTCCCTTGGCGGCACAGGGGACCGCGCGCACCTTGTGGGCTCCAC 558  
Qy 631 GTGCGCGAGGGGACTATGTGATGTATGCTAAATGTGGAAGGCGGTGGTGGCTGAC 690  
Db 559 GTGCGCGGTGGGACTATCTCAGGTTATGCTCAGGCTTGAAGGCTTGTGGTGGGAC 618  
Qy 691 CGGGTTACCTGGAAGGCGCTGATATGTCGGGCAACGCTATTCATCTCCAGTCTTC 750  
Db 619 AGCCCTACCTCCGCGCAGGCTGATGCTGTCGGGCAAGGCGGAGCGGCTTTC 678  
Qy 751 GTGTTTACAGCAACCGTATGGCTGCGGGAACACATTAATGCTTCCGAGGAGAC 810  
Db 679 GTGTCACCAACGCGATGGAGTGTGTAAAGAAACATCGACACTCCAGGCGAT 738  
Qy 811 GTGTTTGGCGGCAATGTTAGGGGTGCGGCAAGGACTTCGGGCTCTCCAC 870  
Db 739 GTGACGTTTGTGGGATGGACAGGAGCTACACCGTGAAGGCTTGGCTCTCACA 798  
Qy 871 CAGTGCACCAACCACTATGACTATTTGGGACCTTGGGATTTGGGCTTCCCTCAAGTC 930  
Db 799 CAGTGCACCAACCACTATGACTATTTGGGACCTTGGGCTTCTGGGCTTCCCTCAAGTC 858  
Qy 931 GGTGTTGATACCACTTACTTAGCCAACTACACCTTCCGATTTCCGTTTCCCTCAAGTC 990  
Db 859 GCGGAGACACTGTCTACCTGCGCAACTTACCTTCCAGACTCTGAGTCTCTGAAGTC 918  
Qy 991 TTAAAGCCAGGAGCGCTTCTTACCCGAAATGGGTGGGCACTCCCTGCGGATCTTCCCA 1050  
Db 919 TTAAAGCCAGGAGCGCTTCTTCCCGGAGTGGGCAATTAATGAGACTTCTCTCCA 978  
Qy 1051 CTC 1053  
Db 979 CTC 981

RESULT 10  
US-08-395-800A-9  
; Sequence 9, Application US/08395800A  
; Patent No. 5867732  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; APPLICANT: LENNON, GREGORY  
; APPLICANT: ROQUIER, SYLVIE  
; APPLICANT: GIORGI, DOMINIQUE  
; APPLICANT: KELLY, ROBERT J.  
; TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE  
; TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE  
; TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF  
; TITLE OF INVENTION: GENOTYPING A PERSON  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: USA  
; ZIP: 22202

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,800A
FILING DATE: 28-FEB-1995
CLASSIFICATION: 435
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1136 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 39..1133
US-08-395-800A-9

Query Match 35.2%; Score 405; DB 1; Length 1136;
Best Local Similarity 68.4%; Pred. No. 2.2e-108;
Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;

QY 217 GGCATGTTACGATCAATTCATTGGCGGCTGGGACACAGATGGGGAATACGCCACA 276
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 273 GGCACCTGGACTGCTACCCCAATGGGCGGTTTGGTAATCAGATGGGACATGCCACG 332
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 277 CTCCTTTCAGCTGGCCAGGATCAAGCGGCTTGGCTTCATCCCGCATCCGATCCATGCACAC 336
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 333 CTGCTGGCTCTGGCCAGCTCAAGGCGCGCGGCCCTTATCCCTGCCATGATGCC 392
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 337 GCTTAGGCGCCATCTTCAGGATCAGGCTCCCGGTGTACACAGCAGACGCGCAAAAG 396
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 393 GCCCTGGCGCGGTTATCCGATCACCCTGCCCTGCTGGCCAGAGAGTGGACGCGC 452
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 397 ATCCCATGGCAGATTACCATCTCAAGCAGCTGGATGGAGGAGGTTACCCGCCATTCGG 456
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 453 ACGCGCTGGCGGGAGCTGCAGCTTCAGACCTGAGATGCGGAGGATAGCGGAGCTTGAGA 512
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 457 GGACACTTTGTGCGCTCACGGGATACCCGTGCTCCCTGACCTTCTACACCACTGCGC 516
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 513 GATCCTTTCGAGCTCTGTGGCTTCCCTGCTCTGGACTTCTTCCACCATCTCGG 572
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 517 CCAGAGATCCGAGAGGTTCAACCTGATGACCACTGCGGAGGAGGCCAGGCGCTTC 576
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 573 GAACAGATCCGACAGAGATTCACCTGTCACACCACTTCGGGAAGAGGCGCAGAGTGTG 632
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 577 CTGCGTGTCTGCGGGTG-----AATGGGAGCAGCGGAGTACTTTGTGGGTGTCAT 630
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 633 CTGGGTAGCTCCCTCGGCGGACAGAGGAGCCGCGCGCCTTGTGCGGCTCCAC 692
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 631 GTGCGCCGAGGGGACTATGTCATCTCATGCTTAATGTGGGAGGGCGTGTGGCTGAC 690
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 693 GTGCGCGTGGGAGCTATCTCAGCTATGCCAGCTATGCCAGAGGTGTGGTGGGGAC 752
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 691 CGGGGTTACCTGGGAAAGCGCCTGGATGATGTCGGGCACTATTCATCTCCAGTCTTC 750
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 753 AGCGCCTACCTCCGCGAGCCATGAGTGTTCGGGCGACGCGCAAGGCCCGGTTTC 812
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 751 GTGGTTACAAGCAACGGTATGCTGCTGGCGGAGACATTAATGCTTCCGAGGAGAC 810
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 813 GTGGTACCAACAGCGGATGAGTGTGTAAGAAACATFCGACACCTCCAGGGGAT 872
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 811 GTGGTGTTCGGGCAATGTTATGAGGGTTCGCCAGCCAGGACTTCGGCTCTCACC 870
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 873 GTGACGTTTGTGCGGATGGACAGAGGCTACACCGTGGAAAGACTTTCCTCTCACA 932
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 871 CAGTGCACCAACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCTACCTGGCA 930
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

933 CAGTGCACCAACCATCATGACTATTGGGACCTTTGGGCTTGGCTTGGCTGCTTGGCTGCTTGGCT 992
931 GGTGGTATACCATCTACTTACCTTACCCAACTACACCCCTTCCGGATTCCTCGTTCCTCAAAGTC 990
993 GCGGAGACACGTCTCTACCTGGCCAACTTACCCCTGCCAGACTCTGAGTTCCTCTGAAGATC 1052
991 TTATAGCCAGAGGACAGCCCTTCTATACCCGAATGGGTGGGCACTCCCTGCGGATCTGTCCCA 1050
1053 TTAAAGCCGAGGCGGCGCTTCTGCGCGAGTGGGTGGGCATTAATGCACTTGTCTCCA 1112
1051 CTC 1053
1113 CTC 1115

RESULT 11
PCT-US95-07554-3
; Sequence 3, Application PC/TUS9507554
; GENERAL INFORMATION:
; APPLICANT: Sandrin, Mauro S.
; APPLICANT: Fodor, William L.
; APPLICANT: Rother, Russell P.
; APPLICANT: Squinto, Stephen P.
; APPLICANT: McKenzie, Ian F. C.
; TITLE OF INVENTION: Methods for Reducing
; TITLE OF INVENTION: Hypersensitive Rejection of Xenografts
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maurice M. Klee
; STREET: 1951 Burr Street
; CITY: Fairfield
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06430
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 750 Kb storage
; COMPUTER: Dell 486/50
; OPERATING SYSTEM: DOS 6.2
; SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07554
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/260,201
; FILING DATE: June 15, 1994
; CLASSIFICATION:
; APPLICATION NUMBER: 08/278,282
; FILING DATE: July 21, 1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Klee, Maurice M.
; REGISTRATION NUMBER: 30,399
; REFERENCE/DOCKET NUMBER: ALX-144.1PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 255-1400
; TELEFAX: (203) 254-1101
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1174 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; MOLECULE TYPE: cDNA to mRNA
; DESCRIPTION: Human H-transferase
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; PUBLICATION INFORMATION:
; AUTHORS: Larsen, R.D.
; AUTHORS: Ernst, L.K.
; AUTHORS: Nair, R.P.
; AUTHORS: Lowe, J.B.
```

```

; TITLE: Molecular cloning, sequence, and
; expression of a human GPP-L-fucose:
; D-galactoside 3-alpha-L-
; fucosyltransferase cDNA that can
; form the H blood group antigen.
; JOURNAL: Proceedings of the National
; Academy of Sciences, USA
; VOLUME: 87
; PAGES: 6674-6678
; DATE: SEP-1990
; PCT-US95-07554-3

```

Query Match	35.2%	Score 405	DB 5	Length 1174
Best Local Similarity	68.4%	Pred. No. 2.3e+08		
Matches 577	Conservative 0	Mismatches 260	Indels 6	Gaps 1
QY	217	GGCATGTTCACGATCAATTCATCTGGCCGCTGGGAAACCAGATGGCGAATACGCCACA	276	
Db	249	GGCACTGGACTGTCTACCCCANATGCCGTTTGGTAATCAGATGGACASTATGCCACG	308	
QY	277	CTCTTTGCACTGGCCAGGATGAACGGACGGCTTTGGTTCTATCCCGCATCCATGCACAAC	336	
Db	309	CTGTGCTCTGGCCAGCTCAACGGCGCGCGGCTTTATCTGCTGCTGCATGCATGCC	368	
QY	337	GCCTAGCGCCCATCTTCAGATCAGCTCCCGTGTATTACAGCGACACGGCCAAAAG	396	
Db	369	GCCTGGCCCGGTATTTCCGCATACCTGCCGCTGTGCCCCAGAAATGGACAGCGCG	428	
QY	397	ATCCCATGGCAGAAATTACCATCTCAACGACTGGATGAGAGAGGTTACCGCCACATTCOG	456	
Db	429	ACGCCGTGGCGAGCTGCAGCTTCAGACTTGGATGTCGGAGGAGTACGGGACTTGAGA	488	
QY	457	GGACACTTTTCGCGCTTCACGGGATACCCGTGCTCTGGACCTTCTACCAACACCTGGCG	516	
Db	489	GAATCTTTCTGAAGCTCTCTGCCTTCCCTGCTCTTGGACTTCTTCCACCATCTCCGG	548	
QY	517	CCAGAGATCTGAAGGAGTTACCCCTGCATGACACAGTGGCGGAGGAGGCCAGGCTTC	576	
Db	549	GAACAGATCGCAGAGAGTTACCCCTGCACGCCACCTTCGGGAAGAGGCGAGAGTGTG	608	
QY	577	CTCGGTGGTTCGGGTG-----AATGGAGCGCCCGAGTACTTTTGGGTGTCCAT	630	
Db	609	CTGGGTAGCTCGCCCTGGCCCGACAGGGACCGCCGCGACCTTTGCGCGCTCCAC	668	
QY	631	GTGCGCGGAGGGGACTATGTSCATGTGTCATGCCCTAATGTGGAAGGCGGTGGTGGCTGAC	690	
Db	669	GTGCGCGTGGGACTATCTGCAGGTATGCTCTCAGCGCTGGAAGSGTGTGGTGGCGAC	728	
QY	691	CGSGGTTACTGGAAGAGCCCTGGATATGTTCCGGCAGCGTATTCATCTCAGTCTTC	750	
Db	729	AGCGCTACCTTCGGGAGGCGCATGGACTGTTTCGGGCGACGGACGAAAGCCCGCTTTC	788	
QY	751	GTGGTTACAAGCAACGGTATGGCTGTGTCGGGGAGAACATTAATGCTCCCGAGGAGAC	810	
Db	789	GTGGTCACCAAGCAACGGCATGGATGGTGTAAAGAAACATTCGACACTCCAGGGCGAT	848	
QY	811	GTGGTGTTCGGGGCAATGTATTTGAGGGTTCGCGAGCGCAAGGACTTCGCGTGTCTACC	870	
Db	849	GTGACGTTTCTGGCATGACAGGAGGCTACACCGTGGAAAGACTTGGCCCTGCTCACA	908	
QY	871	CAGTGCAACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCTACCTGGCA	930	
Db	909	CAGTGCAACCAACCATATGACCATTTGGCACTTCGGCTCTGCGGCTGCTACCTGGCT	968	
QY	931	GGTGGTGATACCATCTACTTAGCCAACTACACCTTCCGGATTTCTCGTTTCCCTCAAGTC	990	
Db	969	GGCGGAGACACTGTCTACCTGGCCAACTTCACCTGCGACACTCTGAGTTCTCTGAAGATC	1028	
QY	991	TTTAAAGCCAGGACCCCTTCCTACCCGATGGGTGGGCATCCCTGCCGATCTGTCCCA	1050	
Db	1029	TTTAAAGCCGAGCGGCTTCCTTCGCGAGTGGGTGGGCATTAATACAGACTTGTCTCCA	1088	
QY	1051	CTC 1053		

Db 1089 CTC 1091

## RESULT 12

```

US-08-395-800A-5
; Sequence 5, Application US/08395800A
; Patent No. 5807732
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B
; APPLICANT: LENNON, GREGORY
; APPLICANT: ROQUIER, SYLVIE
; APPLICANT: GIORGI, DOMINIQUE
; APPLICANT: KELLY, ROBERT J
; TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE
; TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE
; TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF
; TITLE OF INVENTION: GENOTYPING A PERSON
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,800A
; FILING DATE: 28-FEB-1995
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248B85 OPAT UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1199 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 51..1145
; PS-08-395-800A-5

```

	Query Match.	35.2%;	Score 405;	DB 1:	Length 1199;
	Best Local Similarity	68.4%;	Pred. No. 2.3e-108;		
	Matches 577;	Conservative 0;	Mismatches 260;	Indels 6;	Gaps 17;
QY	217	GGGATGTTACGATCAATTC	CAATGGCCGGTGGGGAACCA	GATGGCGGAATACGCCACA	276
Db	285	GGCACCTGGACTGTCTAC	CCCAATGGCGGTTTG	TGTAATCAGATGGACAGIATGCCA	344
QY	277	CTCTTTGCACATGGCCAG	GATGAACGGACGGCTTGC	GGTTTCATCCCGCATCCATGCACAAC	336
Db	345	CTGCGTCTGGCCCAAGCT	CAAGGGCGCGGGGCGCTT	ATCTCCCTGGCATGCATGCC	404
QY	337	GCCTCAGCGCCCATCTTC	CAGGATCAGCCTCCCGTG	TTCACACGCGACAGCGCAAAAAG	396
Db	405	GCCTGSCCCCGTATTC	CGCATCACCTGCCCGTG	CTGSCCCAGAGTGCACAGCCGC	464
QY	397	ATCCATGGCAGAAATTAC	CATCTCAACGACTGGAT	GAGGAGAGCGTTACCGCCCATTCGG	456
Db	465	ACGCGCTGGCGGAGCTG	CAGCTTCACGACTGGAT	TCGGAGGAGIACGGGAGCTTGAGA	524
QY	457	GGACACTTTGTGGCTTC	CAGGGATACCGGTGCT	CTCTGGACCTTCTACCAACCACTCGC	516
b	525	GATCTCTCTCTGAAGCT	CTCTGGCTTCCCTGCT	CTCTTGGACTTCTTCCACCATCTCCGG	584

```
QY 517 COAGAGATCCTGAAGAGAGTTCACCTCATGATGACACAGTCGGGAGGAGGCCAGGCTTC 576
D 585 GAACAGATCCGAGAGAGTTCACCTCATGATGACACAGTCGGGAGGAGGCCAGGAGTGTG 644
QY 577 CTGCGTGTCTCGCGGTG-----AATGGAGCAGCGAGTACCTTTGTTGGTGTCCAT 630
D 645 CTGGGTGAGTCTCGCGGTGCGCGGACAGAGGAGCGCGCGACCTTTGTCGGCGTCCAC 704
QY 631 GTGCGCCGAGGAGGAGTATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 690
D 705 GTGCGCGTGTGGGAGTATCTGAGGTATGCTGAGCGCTGGAGGGGTGTGGTGGGCGAC 764
QY 691 CGGGTTTACCTGGAAAGAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 750
D 765 AGCGCTACCTCCGCGAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 824
QY 751 GTGGTTTCAAGCAAGCGGTATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 810
D 825 GTGGTTTCAAGCAAGCGGTATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 884
QY 811 GTGGTTTTCGGGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 870
D 885 GTGAGGTTTGTGGGAGTGGAGAGGCTACACCGTGGAGAGATTTGCGCTGCTGCTGCTGCTGCT 944
QY 871 CAGTGCACACACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 930
D 945 CAGTGCACACACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1004
QY 931 GTGTGTGATACCATCTACTTACGCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 990
D 1005 GCGGAGACATCTGTACTCGGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1064
QY 991 TTTAAGCCAGAGGAGGCTTCTACCCGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1050
D 1065 TTTAAGCCAGAGGAGGCTTCTACCCGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1124
QY 1051 CTC 1053
D 1125 CTC 1127
```

## RESULT 13

```
US-06-675-773B-4
; Sequence 4, Application US/08675773B
; Patent No. 6166288
; GENERAL INFORMATION:
; APPLICANT: DIAMOND, LISA E
; APPLICANT: LOGAN, JOHN S
; APPLICANT: BYRNE, GUERARD W
; APPLICANT: SHARMA, AJAY
; TITLE OF INVENTION: METHOD OF PRODUCING TRANSGENIC ANIMALS
; TITLE OF INVENTION: FOR XENOTRANSPLANTATION.
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 SEVENTH STREET, N.W., SUITE 300
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,773B
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, IVER P.
; REGISTRATION NUMBER: 28,005
```

```
; REFERENCE/DOCKET NUMBER: DIAMOND-1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2268 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-675-773B-4
```

Query Match 35.2%; Score 405; DB 4; Length 2268;

Best Local Similarity 68.4%; Pred. No. 3.1e-108;

Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;

```
QY 217 GGCATGTTTACAGATCAATTCATTTGCCGSCGTCGGGAAACACAGATGGCGGAATACGCCACA 276
D 749 GGCACCTGGAGTGTCTACCCCAATGGCCGTTTGGTAATCAGATGGGACAGTATGCCAGG 808
QY 277 CTCCTTTCAGTGGCCAGGATGAACGAGGTTTGGTTTCATCCCGGCAATCCATGCAACAAC 336
D 809 CTGCTGGCTCTGGCCAGCTCAACGCGCGCGGCGCTTTATCTTCCCTCCATGATGCC 868
QY 337 GCTCTAGCGCCCATCTTTCAGATCAGCTCCCGGTGTTACAGCGACACAGGCCCAAAAG 396
D 869 GCGCTGGCCCGGATTTCCGATACCTTCCGCTGCTGGCCCAAGAGTGGACAGCGCG 928
QY 397 ATCCCATGGCAGAAATTACCATCTCAACGACTGATGGAGGAGCGTTACGCCACATTCGG 456
D 929 ACGCCGTGGCGGAGCTGACGCTTCACGACTGATGTCGGAGGACTCGGGACTTTGAGA 988
QY 457 GGACACTTTGTGCGCTTACCGGATPACCGTCTCTCTGAGCTTCTACCAACACCTTGGC 516
D 989 GATCCTTTCCCTGAAGCTCTCTGGCTTCCCTGCTCTTGGACTTTCTCCACCATCTCGG 1048
QY 517 CCAGAGATCCCTGAAGAGTTCACCTGATACACAGTGGCGGAGGAGGCCAGGCCCTTC 576
D 1049 GAACAGATCCGACAGAGTTCACCTGACACACACCTTCGGGAGAGGCGGACAGAGTGT 1108
QY 577 CTGGTGTCTGCGGGTG-----AATGGGAGCGGAGCTACTTTTGTGGGTGTCAT 630
D 1109 CTGGTCACTCCGCTGGCGCGCACAGGAGCACCGCCGCGACCTTTGTGCGGTTCAC 1168
QY 631 GTGCGCGAGGAGGACTATGTCATGTCATGCTAATGTGTGGAAGGCGTGTGCTGCTGAC 690
D 1169 GTGCGCGGTGGGAGTATCTGAGGTTATGCTCAGCGCTGGAAGGCTGTGCTGGGCGAC 1228
QY 691 CGGGTTTACCTGGAAAGGCGCTGATGATGTCGGGCGACGCTATTATCTCTCCAGTCTTC 750
D 1229 AGCCCTTACCTCCGCGAGGCGATGACTGTTCCGGGCGAGCGGACGAGCCCGCTTTC 1288
QY 751 GTGTTACAACAGCAACGCTATGCTGCTGGTTCGGGAGAGCAATTAATGCTTCCGAGGAGAC 810
D 1289 GTGTTACACAGCAACGCTATGCTGCTGGTGTGTAAGAAACATCGACACCTCCAGGCGAT 1348
QY 811 GTGTTGTTCGGGCGCAATGTTAGGGGTGTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 870
D 1349 GTGAGCTTGTGGCGATGGACAGGAGCTACACCGTGGAAAGACTTTGCCCTCTCTACA 1408
QY 871 CAGTGCACACACACATCATGACTATTGGGACCTTTGGGATTTGGGATTTGGGATTTGGGATTTGGGAT 930
D 1409 CAGTGCACACACACATCATGACTATTGGGACCTTTGGGATTTGGGATTTGGGATTTGGGATTTGGGAT 1468
QY 931 GGTGGTGATACCATCTACTTAGCCAACTACACCTTCCCGGATTCCTCGCTTCTCCCTCAAGTC 990
D 1469 GCGGAGACACTGTCTACCTGGGCAACTTCACTGCGACACTCTGAGTTCCTCGGAGATC 1528
QY 991 TTTAAGCCAGAGGAGGCTTCTACCGGAATGGTGGGCAATCCCTGCGGATCTGTCGCCCA 1050
D 1529 TTTAAGCCAGAGGAGGCTTCTGCGCGAGTGGTGGGCAATTAATGACAGACTTGTCTCCA 1588
```

```

QY 1051 CTC 1053
Db 1589 CTC 1591

RESULT 14
US-08-273-411-2
; Sequence 2, Application US/08273411
; Patent No. 5625124
; GENERAL INFORMATION:
; APPLICANT: Falk, Per
; APPLICANT: Gordon, Jeffrey I.
; TITLE OF INVENTION: Animal Model for Gastro-Intestinal
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273.411
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 815-6508
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3373 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 104..1201
; OTHER INFORMATION: /note= "Nucleotides 104 through 1201 encode the GDP-L-fuc
; PUBLICATION INFORMATION:
; AUTHORS: Larsen, et al.
; JOURNAL: Proc. Nat'l Acad. Sci. USA
; VOLUME: 87
; PAGES: 6674-6678.
; DATE: 1990
; RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 3373
US-08-273-411-2

Query Match 35.2%; Score 405; DB 1; Length 3373;
Best Local Similarity 68.4%; Pred. No. 3.7e-108;
Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;

QY 217 GGCATGTTCCAGCATCAATTCATTGGCGCGTGGGGAACACAGATGGCGGCAATACGCCACA 276
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 338 GGCACCTGGACTGTCTACCCCAAGCCGGTTGGTAATCAGATGGGACAGTATGCCAG 397

QY 277 CTTCTTGCACTGGCCAGGATGACGAGCGCTTGGCTTCATCCCGCATCCATGCACACAC 336
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 398 CTGCTGGCTCTGCGCCAGCTACAGCGCCGCGGCGCTTATCTCTGCTGCCATGATGCC 457

QY 337 GCTCTAGCGCCCATCTTCAGGATCAGCCCTCCCGGTGTTACACAGGACACGCGCAAAAAG 396

```

```

Db 458 GCCTGGCCCGGTATTCCCGCATCACCTGCCCTGCCCTGGCCCGCAGAGTGSACAGCGCG 517
QY 397 ATCCCATGGCAGAAATTACCATCTCAAGACATGGATGGAGAGCGTTACCGGCCACATCCG 456
Db 518 ACGCCGTGGCGGAGCTGCAGCTTCACGACTGGATGTCGGAGGAGTACCGCGAGTTGAGA 577
QY 457 GGACACTTTGTGCGCTTCAAGGATACCGTGTCTCTGGACCTTCTACACACACTGCGC 516
Db 578 GATCCTTTCTTGAAGCTCTCTGGCTTCCCTGCTCTTGGACTTTCTTCCACCATCTCCGG 637
QY 517 CCAGAGATCTTGAAGGAGTTCCACCTGCATGACACCTTGGCGGAGGAGGCCGCCAGCTTC 576
Db 538 GAACAGATCCCGCAGAGTTCAOCCIGCACACCACTTGGGAAGAGGCGCAGAGTGTG 697
QY 577 CTGCGTGTCTGCGGGTG-----AATGGGAGCAGCGGAGTACTTTTGTGGGTGTCCAT 630
Db 698 CTGGGTCACTCCCGCTGGCGGCGACAGGGGACGCGCGCACCTTTGTGCGGCTCCAC 757
QY 631 GTGCGCGGAGGGGACTATGTGCATCTCATGCTTAATGTGTGAAGGGGCTGTGCTGAC 690
Db 758 GTGCGCGGTGGGACTATCTGAGGTTATGCTCAGCGCTGGAAGGGTGTGGTGGGCGAC 817
QY 691 CGGGGTTACCTGGAAAGGCCCTGCATATGTTCGGGSCACGCTATTCACTCCASTCTTC 750
Db 818 AGCGCTACCTCCGCGAGGCCATGCTGTTCCGGSCACGCGCACGAAGCCCGCTTTC 877
QY 751 GTGTTACAAGCAACGGTATGCGCTGTGCGGAGAGACATTAATGCTTCCCGGAGGAC 810
Db 878 GTGGTCAACCAACGGCATGAGTGTGTAAAGAAACATCGACACCTCCGAGGGCGAT 937
QY 811 GTGGTGTTCGGGCAATGGTATTGAGGGTGCAGCGGCTCCGCAAGGACTTCGCGCTGCTACC 870
Db 938 GTGAGCTTTGCTGGGATGGAGAGGCTACACGCTGGAAGAGCTTTGCCCTCTCACA 997
QY 871 CAGTGCAACCAACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCGCTACCTGGCA 930
Db 998 CAGTGCAACCAACCATCATGACTATTGGGACCTTTGGGCTGCGCTGCGCTGCGCTGCGCT 1057
QY 931 GGTGTGATACCATCTACTTACGCACTACACCTTCCGGATTCCTCCCTCAAGTGC 990
Db 1058 GCGGAGAGACTGTCTACCTGGCCAACTTCCGCTGCGGAGCTCTGAGTTCTTGAAGTCT 1117
QY 991 TTAAAGCCAGAGGCGCGCTTCTACCCGAATGGTGGGATCCCTGCGGATCTGTCCCA 1050
Db 1118 TTAAAGCCAGGCGCGCTTCTGCGGAGTGGGCTTAAAGCAGACTTGTCTTCTCCA 1177
QY 1051 CTC 1053
Db 1178 CTC 1180

RESULT 15
US-08-675-773B-3
; Sequence 3, Application US/08675773B
; Patent No. 6156288
; GENERAL INFORMATION:
; APPLICANT: DIAMOND, LISA E
; APPLICANT: LOGAN, JOHN S
; APPLICANT: BYRNE, GUERARD W
; APPLICANT: SHARMA, AJAY
; TITLE OF INVENTION: METHOD OF PRODUCING TRANSGENIC ANIMALS
; TITLE OF INVENTION: FOR XENOTRANSPLANTATION.
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 SEVENTH STREET, N.W., SUITE 300
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/675,773B  
FILING DATE: 03-JUL-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, IVER P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: DIAMOND=1A  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3791 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-575-773B-3

Query Match 35.2%; Score 405; DB 4; Length 3791;  
Best Local Similarity 68.4%; Pred. No. 3.9e-108;  
Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;

217 GGCATGTTACAGATCAATTCATTCGCGCTGGGAAACAGATGGCGAATAGCCACACA 276  
2293 GGACCTGGACTGTCFACCCCAATGGCGGTTGGTAATCAGATGGGACATATGCCAAG 2352  
277 CTCTTTGCATGGCCAGGATGACGGAGCGTTGCGTTTCATCCCGCATCCATGCAAC 336  
2353 CTGCTGCTCTGGCCCAAGCTCAACGGCGCGCGGCTTATCTGCTGCTGCATGCTGCC 2412  
337 GCTCTAGCGCCATCTTCAGATCAGCTCCCGCTGTTTACACAGCGACAGGCCAAAAG 396  
2413 GCCCTGGCCCCGGTATTCGGCATCACCCTGCCCTGCTGGCCCCAGAGTGGACAGCGC 2472  
397 ATCCCATGGCAGAAATACCAATCTCAAGACTGATGGAGAGCGTTACCGCCACATCCG 456  
2473 ACGCCGTGGCGGAGCTGCACCTTCAGACTGGATGTCGGAGAGGATACGGGACTTGAGA 2532  
457 GGACACTTTGTGGCGCTTCAAGGATACCCGTCTCTGGACCTTACACACACCTGGGC 516  
2533 CATCTTTCTTGAAGCTCTCTGGCTCCCGCTCTTGGACTTCTTCCACCATCTCCGG 2592  
517 CCAGAGATCTTGAAGGATTCACCTCATGACACCTGCGGGAGGAGGCCAGSCCTTC 576  
2593 GAACAGATCCCGAGAGATTCACCTTGACGACACCTTCGGGAAGAGGCGCAGAGTGTG 2652  
577 CTGCTGCTCTGCGGGTG-----AATGGAGCCAGCCGAGTACTTTTGGGTGTCAT 630  
2653 CTGGGTGAGCTCCGCTGGCGCGACAGGGAGCCCGCGCACCTTTGTCGGCGTCCAC 2712  
631 GTGCGCCGAGGGGATGATGTCATGCTATGCTATGTTGGAAGGGGTGTTGGCTGAC 690  
2713 GTGCGCGCTGGGGACTATCTGAGGTTATGCTCATGCGCTGGAAGGGTGTGGCGCAC 2772  
691 CGGGGTACTCTGMAAAGCCCTGGATATGTCGGGGCACGCTATTCATCTCCAGTCTTC 750  
2773 AGCGCCCTACTCTCGGCGAGGCGATGGACTGTGTCGGGCGACGCGACAGCCCGCTTC 2832  
751 GTGTTTACAAGCAACCGATGCGCTGGTCGGGAGAAATTAATGCTTCCCGAGGAGAC 810  
2833 GTGGTACCAGCAACGCGATGGATGTTGTAAGAAACATCGACACCTCCAGGGCGAT 2892  
811 GTGGTGTCCCGGGCAATGATGAGGGTCCCGACCAAGCACTTCGGCTGCTCAAC 870  
2893 GTGACGTTTCTTGGCGATGACAGAGGCTTACACCGTGGAAAGACITTCGCTGCTACA 2952  
871 CAGTGAACCAACCATCATGACTATGGACCTTTGGGATTTGGGCTGCTACCTGGCA 930

2953 CAGTCAACCAACCACTATGACCAITGGCACTTCGGCTTCCTGGCTGCCCTACCTGGCT 3012  
931 GGTGGTGATACATCTACTTAGCCAACTACACCTTCGGGATCTCCGTTTCCTCAAGTC 990  
3013 GCGGAGACACTGTCTACTCTGGCCAACTTCACCTGCCAGACTCTGAGTCTCTGAAGATC 3072  
991 TTTAAGCCAGAGGCGAGCTTCTTACCCGGAAGGGTGGGATCCCTGCGGATCTGTCCCA 1050  
3073 TTTAAGCCGAGGCGGCTTCTGCGCCGAGTGGGTGGGATTAATGACAGACTGTCTCCA 3132  
1051 CTC 1053  
3133 CTC 3135

Search completed: May 26, 2003, 11:58:32  
Job time : 64.7104 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

      OM nucleic - nucleic search, using sw model

Run on:      May 26, 2003, 11:58:37 ; Search time 176.211 Seconds
              (without alignments)
              8610.191 Million cell updates/sec

Title:
Perfect score: 1149
Sequence: 1 atgccagcgccacggttcc.....gaagagcctctgatggaa ll49

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:      828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters:      1657494

Minimum DB seq length: 0
Maximum DB seq length: 200000000

```

PRRESLT  
US-09-  
; Sequen  
; Pater  
; GENEE  
APPL  
TITT  
TITT  
TITT  
FILL  
CURR  
CURR  
PRII  
PRII

1	1149	100.0	1149	10	US-09-999-672-7	Sequence 7, Appl
2	1149	100.0	1149	12	US-10-040-863-7	Sequence 7, Appl
3	1068	93.0	1068	10	US-09-999-672-9	Sequence 9, Appl
4	1068	93.0	1068	12	US-10-040-863-9	Sequence 9, Appl
5	630.2	54.8	1043	10	US-09-051-034A-1	Sequence 1, Appl
6	632.6	53.3	3088	10	US-09-954-456-45	Sequence 45, Appl
7	612.6	53.3	3088	10	US-09-954-456-1621	Sequence 1621, Ap
8	612.6	53.3	3088	10	US-09-969-347-234	Sequence 234, App
9	423.8	36.9	1098	10	US-09-051-034A-3	Sequence 3, Appl
10	423.8	36.9	1269	10	US-09-844-268-12	Sequence 12, Appl
11	423.8	36.9	1269	10	US-09-844-705-12	Sequence 12, Appl
12	405	35.2	3373	9	US-10-105-963-9	Sequence 9, Appl
13	405	35.2	8174	10	US-09-863-475A-5	Sequence 5, Appl
14	72.8	6.3	100	10	US-09-999-673-29	Sequence 29, Appl
15	72.8	6.3	100	12	US-10-040-863-29	Sequence 29, Appl
16	71.2	6.2	100	10	US-09-999-672-23	Sequence 23, Appl
17	71.2	6.2	100	10	US-09-999-673-26	Sequence 26, Appl
18	71.2	6.2	100	12	US-10-040-863-23	Sequence 23, Appl
19	71.2	6.2	100	12	US-10-040-863-26	Sequence 26, Appl

181 CGGACAGCAGCAGCATGGAAATGGAGAGCTGGGGCAATGTCACGATCAATTCATT 240  
241 GGCGCGTGGGACCAAGATGGGCGATAGCCACACTTTTCACHTGGCCAGGATGAAC 300  
241 GGCGCGTGGGACCAAGATGGGCGATAGCCACACTTTTCACHTGGCCAGGATGAAC 300  
301 GGAGGGTTGGTTCACTCCCGCATCCATGACAAACGCTCTAGGCGCCACTTTTCAGGATC 360  
301 GGAGGGTTGGTTCACTCCCGCATCCATGACAAACGCTCTAGGCGCCACTTTTCAGGATC 360  
361 AGCCTCCCGGTTTACACAGCGACAGCCGCAAAAGATGCCATGGCAGATTCATCTC 420  
361 AGCCTCCCGGTTTACACAGCGACAGCCGCAAAAGATGCCATGGCAGATTCATCTC 420  
421 AACGACTGGATGGAGGAGCGTTACGCGACACTTCGCGACACTTTGTCGCTTCACGGGA 480  
421 AACGACTGGATGGAGGAGCGTTACGCGACACTTCGCGACACTTTGTCGCTTCACGGGA 480  
481 TACCGGTCTCTGGACCTTCTACACACACTGCGCGCCAGAGATCCTGAAGAGTTACCC 540  
481 TACCGGTCTCTGGACCTTCTACACACACTGCGCGCCAGAGATCCTGAAGAGTTACCC 540  
541 CTGCATGACACAGCTGGGGAGAGGCCGCTTCCTGCTGGTCTCTGCGGTGAATGGG 600  
541 CTGCATGACACAGCTGGGGAGAGGCCGCTTCCTGCTGGTCTCTGCGGTGAATGGG 600  
601 AGCCAGCGAGTACTTTTCTGGGTCTCATGTGCGCGAGGACTATGTCATGTCATG 660  
601 AGCCAGCGAGTACTTTTCTGGGTCTCATGTGCGCGAGGACTATGTCATGTCATG 660  
661 CTAATGTTGGAGGGGCTGTGCTGACCGGGTTACCTGGAAAGGCCCTTGGATATG 720  
661 CTAATGTTGGAGGGGCTGTGCTGACCGGGTTACCTGGAAAGGCCCTTGGATATG 720  
721 TTCGGGCAAGCTATATCTCCATCTCCAGTCTCGTGGTTACAAAGCAAGGATGGCTG 780  
721 TTCGGGCAAGCTATATCTCCATCTCCAGTCTCGTGGTTACAAAGCAAGGATGGCTG 780  
781 CGGAGACACTTAATGCTCCCGAGAGAGCGTGGTGGTTCGCGGCAATGGATTGAGGG 840  
781 CGGAGACACTTAATGCTCCCGAGAGAGCGTGGTGGTTCGCGGCAATGGATTGAGGG 840  
841 TCGCCAGCAAGGACTTCGCGTGTCTACCCAGTGCACACACACCATCATGACTATTGG 900  
841 TCGCCAGCAAGGACTTCGCGTGTCTACCCAGTGCACACACACCATCATGACTATTGG 900  
901 ACCTTTGGGATTTGGGCTCCCTACCTGGCAGTGGTGATACCATCTACTTACCCACTAC 960  
901 ACCTTTGGGATTTGGGCTCCCTACCTGGCAGTGGTGATACCATCTACTTACCCACTAC 960  
961 ACCCTTCGGATTTCCGTTCTCAAAGTCTTTAAGCCAGAGGCGCTTCCTACCGGAA 1020  
961 ACCCTTCGGATTTCCGTTCTCAAAGTCTTTAAGCCAGAGGCGCTTCCTACCGGAA 1020  
1021 TGGGTGGGCATCCCTGCCGATCTGTCGCCACTCTTAAAGGATTAACACAGCGCTGCT 1080  
1021 TGGGTGGGCATCCCTGCCGATCTGTCGCCACTCTTAAAGGATTAACACAGCGCTGCT 1080  
1081 CGGTCCCATTCACCTCAAGCAAAAGAGTACCTTTGTTACGTCGAGAGAGCGCTTC 1140  
1081 CGGTCCCATTCACCTCAAGCAAAAGAGTACCTTTGTTACGTCGAGAGAGCGCTTC 1140  
1141 TGATGGGAA 1149  
1141 TGATGGGAA 1149

RESULT 2  
US-10-040-863-7  
; Sequence 7, Application US/10040863  
; Patent No. US20020137165A1  
; GENERAL INFORMATION:

APPLICANT: Eric H. Holmes et al.  
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
TITLE OF INVENTION: GMI-SPECIFIC ALPHA-2 FUCOSYLTRANSFERASE AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: 8511-029  
CURRENT APPLICATION NUMBER: US/10/040,863  
CURRENT FILING DATE: 2001-11-01  
PRIOR APPLICATION NUMBER: 09/299,886  
PRIOR FILING DATE: 1999-04-23  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 7  
LENGTH: 1149  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1143)  
US-10-040-863-7  
Query Match 100.0%; Score 1149; DB 12; Length 1149;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AIGCCACGCGCCAGGTTCCCTTCCTTCCTCTGCGCCACTTCCTCATCTTCTCTTC 60  
DB 1 ATGCGCAGCGCCAGGTTCCCTTCCTTCCTCTGCGCCACTTCCTCATCTTCTCTTC 60  
QY 61 GTGACTTCCACCATCATCCACTCCAGCAGCGAATAGTGAAGCTCCAAACCCCTGTCAAG 120  
DB 61 GTGACTTCCACCATCATCCACTCCAGCAGCGAATAGTGAAGCTCCAAACCCCTGTCAAG 120  
QY 121 AAGCAATTTACCGATCAGCTCAATATGCTCTGGGAAACACAGAAAGCCAGAGTGGGA 180  
DB 121 AAGCAATTTACCGATCAGCTCAATATGCTCTGGGAAACACAGAAAGCCAGAGTGGGA 180  
QY 181 CGGACAGCGACGACGATGGGAATGAGAGCTGCGGGCATCTTCACGATCAATTCATT 240  
DB 181 CGGACAGCGACGACGATGGGAATGAGAGCTGCGGGCATCTTCACGATCAATTCATT 240  
QY 241 GCGCGCTGGGAAACAGATGGCGCAATACGCCACACTTTTGCATGSCCAGGATGAAC 300  
DB 241 GCGCGCTGGGAAACAGATGGCGCAATACGCCACACTTTTGCATGSCCAGGATGAAC 300  
QY 301 GAGCGCTTGGTTCATCCCGCATCCATGCAACAGCTCTAGCGCCCATCTTCAGGATC 360  
DB 301 GAGCGCTTGGTTCATCCCGCATCCATGCAACAGCTCTAGCGCCCATCTTCAGGATC 360  
QY 361 AGCCTCCCGGTGTTACACAGCGACAGCGCAAAAGATCCCATGGCAGAAATACCATCTC 420  
DB 361 AGCCTCCCGGTGTTACACAGCGACAGCGCAAAAGATCCCATGGCAGAAATACCATCTC 420  
QY 421 AACGACTGGATGGAGGAGCGTTACCCGCACATTCGCGGACACTTTGTGCGCTTCAGGGA 480  
DB 421 AACGACTGGATGGAGGAGCGTTACCCGCACATTCGCGGACACTTTGTGCGCTTCAGGGA 480  
QY 481 TACCGGTCTCTGGACCTTCTACACACACTGCGCGCCAGAGATCCTGAAGAGTTACCC 540  
DB 481 TACCGGTCTCTGGACCTTCTACACACACTGCGCGCCAGAGATCCTGAAGAGTTACCC 540  
QY 541 CTGCATGACACGCTGGGAGGAGCGCCAGGCTTCCTGCTGGTCTGCGGTGAATGGG 600  
DB 541 CTGCATGACACGCTGGGAGGAGCGCCAGGCTTCCTGCTGGTCTGCGGTGAATGGG 600  
QY 601 AGCCAGCGAGTACTTTTGTGGGTGTCATGTCGCGCGAGGAGTATGTCATGTCATG 660  
DB 601 AGCCAGCGAGTACTTTTGTGGGTGTCATGTCGCGCGAGGAGTATGTCATGTCATG 660  
QY 661 CTAATGTTGGAGGGGCTGTGCTGACCGGGTTACCTGGAAAGGCCCTTGGATATG 720  
DB 661 CTAATGTTGGAGGGGCTGTGCTGACCGGGTTACCTGGAAAGGCCCTTGGATATG 720  
QY 721 TTCGGGCAAGCTATATCTCCATCTCCAGTCTCGTGGTTACAAAGCAAGGATGGCTG 780

```
Db 721 TCCGGGACACGTAATCATCTCCAGCTCTCGTGGTTACAGCAAGGATGATGCTGGTGC 780
QY 781 CGGAGAACATTAATGCTTCCCGAGGAGACGTGGTTCGGGGCAATGGTAATGAGGG 840
Db 781 CGGAGAACATTAATGCTTCCCGAGGAGACGTGGTTCGGGGCAATGGTAATGAGGG 840
QY 841 TCGCCAGCAAGGACTTCGGCTGTCTCCAGTGCACACCAACCAATCACTACTATGG 900
Db 841 TCGCCAGCAAGGACTTCGGCTGTCTCCAGTGCACACCAACCAATCACTACTATGG 900
QY 901 ACCTTTGGATTTGGCTGCTACCTGCGAGTGGTGTATACCATCTACTTACCACTAC 960
Db 901 ACCTTTGGATTTGGCTGCTACCTGCGAGTGGTGTATACCATCTACTTACCACTAC 960
QY 961 ACCCTTCGGGATTCCTCGCTTCCCAAGTCTTAAAGCCAGAGCGCTTCTACCCGAA 1020
Db 961 ACCCTTCGGGATTCCTCGCTTCCCAAGTCTTAAAGCCAGAGCGCTTCTACCCGAA 1020
QY 1021 TGGGTGGGATCCCTGCGGATCTGCCCCACCTCTTAAGGCAATTAACACCACTGCTCT 1080
Db 1021 TGGGTGGGATCCCTGCGGATCTGCCCCACCTCTTAAGGCAATTAACACCACTGCTCT 1080
QY 1081 CGGTCCCACTTCCACTCAAGCAAGGAGTCACTTGTACGTGCGAGGAGAGGCTTC 1140
Db 1081 CGGTCCCACTTCCACTCAAGCAAGGAGTCACTTGTACGTGCGAGGAGAGGCTTC 1140
QY 1141 TGATGGGAA 1149
Db 1141 TGATGGGAA 1149
```

## RESULT 3

```
US-09-999-672-9
; Sequence 9, Application US/09999672
; Patent No. US20020127655A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/999,672
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US/09/298,886
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1068
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1062)
US-09-999-672-9
```

```
Query Match 93.0%; Score 1068; DB 10; Length 1068;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 82 CTCACAGCAGGATAGTGAAGCTCCACACCCCTGTGCAGAGGAGGATTAACCGATCAGCACT 141
Db 1 CTCACAGCAGGAGATAGTGAAGCTCCACACCCCTGTGCAGAGGAGGATTAACCGATCAGCACT 60
QY 142 CAATGTCTCGGGAACACAGAAAGCCAGAGATGCGACGGGACAGCAGCAGCATGGG 201
Db 61 CAATGTCTCGGGAACACAGAAAGCCAGAGATGCGACGGGACAGCAGCAGCATGGG 120
QY 202 AATGAGAGCTGCGGGGATGTTCCAGATCAATTCATTTGGCGGCTGGGGAACCATGAT 261
Db 121 AATGAGAGCTGCGGGGATGTTCCAGATCAATTCATTTGGCGGCTGGGGAACCATGAT 180
```

```
QY 262 GCGCAATAGCCACACTCTTTGCACTGGCCAGGATCAAGGAGGCTTGCCTTCATCCCC 321
Db 181 GCGCAATAGCCACACTCTTTGCACTGGCCAGGATCAAGGAGGCTTGCCTTCATCCCC 240
QY 322 GCATCCATGCACAAAGCTCTAGCGCCCATCTTCAGGATCAGCTCCCGGTGTATACAGC 381
Db 241 GCATCCATGCACAAAGCTCTAGCGCCCATCTTCAGGATCAGCTCCCGGTGTATACAGC 300
QY 382 GACAGGGCCAAAAGATCCCATGGGAGAAATACCATCTCAAGACTGGATGGAGGAGCGT 441
Db 301 GACAGGGCCAAAAGATCCCATGGGAGAAATACCATCTCAAGACTGGATGGAGGAGCGT 360
QY 442 TACCCGCCACATTCGGGACACTTTGCGCTTCAAGGATACCGGCTCTCCGACCTTC 501
Db 361 TACCCGCCACATTCGGGACACTTTGCGCTTCAAGGATACCGGCTCTCCGACCTTC 420
QY 502 TACCACCACTGCGGCCAGAGATCTCTGAAGGAGTTACCCCTGCATGACACGTCGGGAG 561
Db 421 TACCACCACTGCGGCCAGAGATCTCTGAAGGAGTTACCCCTGCATGACACGTCGGGAG 480
QY 562 GAGGCCAGGCTTCCCTGCTGCTGGGGTGAATGGGAGCCAGCGGAGTCTTTGTG 621
Db 481 GAGGCCAGGCTTCCCTGCTGCTGGGGTGAATGGGAGCCAGCGGAGTCTTTGTG 540
QY 622 GGTGTCCATGTGCGCGAGGGGACTATGTGCTCATGCTAATGTGTGAAGGCGGTG 681
Db 541 GGTGTCCATGTGCGCGAGGGGACTATGTGCTCATGCTAATGTGTGAAGGCGGTG 600
QY 682 GTGCTGACCGGGGTTACCTGSHAAAGGCCCTGGATATGTCGGGACGCTATTCATCT 741
Db 601 GTGCTGACCGGGGTTACCTGSHAAAGGCCCTGGATATGTCGGGACGCTATTCATCT 660
QY 742 CCAGTCTTCTGCTTACAGCAAGCGTATGCTGCTGCTGCGGAGGACATTAATGCTTC 801
Db 661 CCAGTCTTCTGCTTACAGCAAGCGTATGCTGCTGCTGCGGAGGACATTAATGCTTC 720
QY 802 CGAGGAGACGTGTGTTCGGGGCAATGTTAGGGGTGCGCAGCCAGCAAGCACTTCGG 861
Db 721 CGAGGAGACGTGTGTTCGGGGCAATGTTAGGGGTGCGCAGCCAGCAAGCACTTCGG 780
QY 862 CTGCTCACCCAGTGCACACCAACCATCATGACATTTGGGACCTTTGGGATTTGGCTGCC 921
Db 781 CTGCTCACCCAGTGCACACCAACCATCATGACATTTGGGACCTTTGGGATTTGGCTGCC 840
QY 922 TACCTGCGAGTGGTGTATACCATCTACTTACCACTTACCACTTCCGCTTCCTGCTTC 981
Db 841 TACCTGCGAGTGGTGTATACCATCTACTTACCACTTACCACTTCCGCTTCCTGCTTC 900
QY 982 CTCAGAGTCTTTAAGCCAGAGGAGCGCTTCTACCCGGAATGGGTGGCATCCCTGCGCAT 1041
Db 901 CTCAGAGTCTTTAAGCCAGAGGAGCGCTTCTACCCGGAATGGGTGGCATCCCTGCGCAT 960
QY 1042 CTGTCCTCCACTCCTTAAGGATTAACACCAAGCTTCTGCTGCTGCTCCACTTCCACTCAAG 1101
Db 961 CTGTCCTCCACTCCTTAAGGATTAACACCAAGCTTCTGCTGCTGCTCCACTTCCACTCAAG 1020
QY 1102 GCAAGAGGATCACTTGTAGTGCAGAGGAGCGCTTCTGATGGGA 1149
Db 1021 GCAAGAGGATCACTTGTAGTGCAGAGGAGCGCTTCTGATGGGA 1068
```

## RESULT 4

```
US-10-040-863-9
; Sequence 9, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; CURRENT FILING DATE: 2001-11-01
```

```
; PRIOR APPLICATION NUMBER: 09/298,886
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1068
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1062)
US-10-040-863-9

Query Match      93.0%; Score 1068; DB 12; Length 1068;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  82 CTCAGCAGCGAATAGTGAAGCTCCACCCCTGTGAGAGAGGAATTAACGATGACGACT 141
Db  1 CTCAGCAGCGAATAGTGAAGCTCCACCCCTGTGAGAGAGGAATTAACGATGACGACT 60

Qy  142 CAATGTCCTCGGGAACACAGAAAGCCAGAGATGCGAGCGGACAGGACGATGGG 201
Db  61 CAATGTCCTCGGGAACACAGAAAGCCAGAGATGCGAGCGGACAGGACGATGGG 120

Qy  202 AATGAGAGCTGCGGGGCATGTTTCAAGATCAATTCATTTGGCCGGCTGGGAACCAAG 261
Db  121 AATGAGAGCTGCGGGGCATGTTTCAAGATCAATTCATTTGGCCGGCTGGGAACCAAG 180

Qy  262 GCGAATATGCGCACACTCTTTGCACTGCGCAGATGAACGAGCGGTGTGCTCATGCC 321
Db  181 GCGAATATGCGCACACTCTTTGCACTGCGCAGATGAACGAGCGGTGTGCTCATGCC 240

Qy  322 GCATCCATGACACAGCTCTAGCCGCCATCTCAGGATCAGCCTCCGCGTTTACACAG 381
Db  241 GCATCCATGACACAGCTCTAGCCGCCATCTCAGGATCAGCCTCCGCGTTTACACAG 300

Qy  382 GACACGGCCAAAAGATCCCATGTCAGAAATTAACATCTCAACGACTGGATGGAGAGCGT 441
Db  301 GACACGGCCAAAAGATCCCATGTCAGAAATTAACATCTCAACGACTGGATGGAGAGCGT 360

Qy  442 TACGCGCACATTCGCGGACACTTTGTGGCTTACGGGATACCGGTGCTCGTACGCTTC 501
Db  361 TACGCGCACATTCGCGGACACTTTGTGGCTTACGGGATACCGGTGCTCGTACGCTTC 420

Qy  502 TACCACACCTGCGCCGACAGATCTGGAAGAGTTCACCTGATGACACGTCGCGGAG 561
Db  421 TACCACACCTGCGCCGACAGATCTGGAAGAGTTCACCTGATGACACGTCGCGGAG 480

Qy  562 GAGGCCAGGCGCTTCCTGCGTGGTCTGCGGTGAATGGAGCCAGCCGAGTACTTTTGTG 621
Db  481 GAGGCCAGGCGCTTCCTGCGTGGTCTGCGGTGAATGGAGCCAGCCGAGTACTTTTGTG 540

Qy  622 GGTGTTCATGTGCGCCGAGGAGTATGTGATGATGATGATGATGATGATGATGATGATG 681
Db  541 GGTGTTCATGTGCGCCGAGGAGTATGTGATGATGATGATGATGATGATGATGATGATG 600

Qy  682 GTGGTGACCGGGTTACTTGGAAAAGCCCTGGATATGTTCCGGGACGCTATTCACTCT 741
Db  601 GTGGTGACCGGGTTACTTGGAAAAGCCCTGGATATGTTCCGGGACGCTATTCACTCT 660

Qy  742 CCAGTCTTCGTGGTTTACAGCAACGGTATGCGCTGGTGGCCGGGAGAACATTAATGCTTCC 801
Db  661 CCAGTCTTCGTGGTTTACAGCAACGGTATGCGCTGGTGGCCGGGAGAACATTAATGCTTCC 720

Qy  802 CGAGGAGAGCTGGTGTTCGCGGGAATGGTATTGAGGGGTGCGCCAGGACCTTCGCG 861
Db  721 CGAGGAGAGCTGGTGTTCGCGGGAATGGTATTGAGGGGTGCGCCAGGACCTTCGCG 780

Qy  862 CTGCTCACCCAGTGAACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCC 921
Db  781 CTGCTCACCCAGTGAACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCC 840

; PRIOR APPLICATION NUMBER: 09/051,034A-1
; Sequence 1, Application US/09051034A
; Patent No. US20010055584A1
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, IAN FARQUHAR CAMPBELL
; APPLICANT: SANDRIN, MAURO SERGIO
; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC
; FILE REFERENCE: 30562.6USWO
; CURRENT APPLICATION NUMBER: US/09/051.034A
; CURRENT FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: PCT/AU97/00492
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: 60/024,279
; PRIOR FILING DATE: 1996-08-21
; PRIOR APPLICATION NUMBER: P01402
; PRIOR FILING DATE: 1996-08-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1043
; TYPE: DNA
; ORGANISM: Sus Domesticus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (9)..(1031)
US-09-051-034A-1

Query Match      54.8%; Score 630.2; DB 10; Length 1043;
Best Local Similarity 76.9%; Pred. No. 2.2e-191;
Matches 814; Conservative 0; Mismatches 203; Indels 42; Gaps 2;

Qy  1 ATGCCACGCGCCGAGGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
Db  9 ATGCTCAGCATGCGAGGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 68

Qy  61 GTGACTTCCACCATATCCACCTCCAGCAGGATAGTGAAGCTCCACACCCCTGTGACAG 120
Db  69 ACGGCTTCCACCATATTTCACTTTCAGCAGAGGATGTTGAAGATTCACCCACCT----- 123

Qy  121 AAGGAATTTACGATGACGACTCAAAATGCTCTGGGAACACAGAAAGCCAGAGATGCGA 180
Db  124 -GGGAGTTACAGATGGTGACCGAGTGACACAGAGACCCCTCGAGCCC----- 172

Qy  181 CGGACACGCGAGCAGCATGGAATGGAGAGCTGGGGGSCATGTTCAAGATCAATTCGAT 240
Db  173 -----CCAGCTGAAGGSCATGTGGACCATCAATGCGCATC 206

Qy  241 GCGCGGCTGGGGAACACAGATGGGGAATACCCACACTCTTTGACACTGGCCAGGATGAAC 300
Db  207 GCGCGCTGGGGAACACAGATGGGGAATACCCACCTGTACGGGCTGCCAGGATGAAC 266

Qy  301 GGAGGGCTTGGTTTCATCCCGCCATCCATGACACACGCTCTAGCGCCCACTCTTCAGGATC 360
Db  267 GGGGGGCGGGCTTCATCCCGCCCGAGATGACACAGCGCTGGGCCCCCATCTTCAGGATC 326
```





QY	729	ACGCTAATTCATCCACGTCCTCGTGGTTACAAAGCAACGGTATGGCCTGGTCCGGGAGAA	789
Db	807	TCGCTACAGCTTCCTCACTTCGTTGGTCACCAAGTAATGGCATGGCCTGGTTCGGAGAA	866
QY	789	CATTAATGCTTCCGGAGGACAGCTGGTTCGGGGCAATGGTATTGAGGGGTGCCGACG	848
Db	867	CATTGACACCTCCACGGTGATGTGGTGTGGCGATGGCATTTGAGGCTCACTGC	926
QY	849	CAAGGACTTCGGGTGCTCAACCAGTGCACACCAACCATCATGACTATTGGGACCTTTGG	908
Db	927	CAAGATTTTGGTCTACTCACAGTGAACCAACCATCATGACCAATTTGGAGCTTCGG	986
QY	909	GATTTGGGCTGCTACTGGCAGGTGGTGATACCATCTACTTATGCCAACTACACCTTCC	968
Db	987	GATCTGGGCGCATACCTCAAGGCGGAGACACCATCTACTGCGCAATTTACACCTTCC	1046
QY	969	GGATTTCGCTTCTCAAGTCTTTAAGCCAGAGGAGCCTTCTTATCCCAATTTGGTGGG	1028
Db	1047	CGACTGCCCTTCTCTCAAAATCTTTAAGCCAGAGGAGCCTTCTTGGCGAGTGGACAG	1106
QY	1029	CATCCCTGCGGACTGTCGCCACTCTCTTAAGGCAATTAACACGACCTGCTCTGGTCCCA	1088
Db	1107	GATTGGCGAGACTGTCCCTCTACTCAGACACTAATGCTGGCCCAATCTTTGAGACCT	1166
QY	1089	CTTCCACCTCAAGGC	1103
Db	1167	TTTCTGCTCTCTGC	1181
RESULT 9			
US-09-051-034A-3			
; Sequence 3, Application US/09051034A			
; Patent No. US20010055584A1			
; GENERAL INFORMATION:			
; APPLICANT: MCKENZIE, IAN FARQUHAR CAMPBELL			
; APPLICANT: SANDRIN, MAURO SERGIO			
; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC			
; TITLE OF INVENTION: GLYCOSYLTRANSFERASE			
; FILE REFERENCE: 30562.6USNO			
; CURRENT APPLICATION NUMBER: US/09/051,034A			
; CURRENT FILING DATE: 1998-03-31			
; PRIOR APPLICATION NUMBER: PCT/AU97/00492			
; PRIOR FILING DATE: 1997-08-01			
; PRIOR APPLICATION NUMBER: 60/024,279			
; PRIOR FILING DATE: 1996-08-21			
; PRIOR APPLICATION NUMBER: P01402			
; PRIOR FILING DATE: 1996-08-02			
; NUMBER OF SEQ ID NOS: 14			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 3			
; LENGTH: 1098			
; TYPE: DNA			
; ORGANISM: Sus Domesticus			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1)..(1098)			
US-09-051-034A-3			
Query Match. 36.9%; Score 423.8; DB 10; Length 1095;			
Best Local Similarity 59.6%; Pred. No. 2.6e-125;			
Matches 591; Conservative 0; Mismatches 252; Indels 6; Gaps 1;			
QY	217	GGCATGTTCACGATCAATTCATTGGCCGGCTGGGAAACCATGATGGCGAATACGCCACA	276
Db	235	GGGACCTGGACTATTTCACCGGATGGCCGGTTTGGGAACCATGATGGACAGTATGCCACG	294
QY	277	CTCTTTTGCACTGGCCAGGATGAACGGACGGCTTTCGCTTCATCCCGCCATCCATGCACAAC	336
Db	295	CTGCTGGCCCTGGCGGAGCTCAACGGCGCCAGGCCCTTCATCCAGCCTCCCATGACGCC	354
QY	337	GCNCTAGCGCCCATCTTTCAGGATPACGCTCCCGGGTTTACACAGGACACGGCCAAAAG	396
Db	355	GTCTGGCCCGCGTGTTCGATCATACGCTGCCTGTCTGGCGCCGAGGTAGACAGGCAC	414





```

: APPLICANT: VOGELI, PETER
: TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY
: TITLE OF INVENTION: RESISTANT TO F18 E. COLI ASSOCIATED DISEASES.
: FILE REFERENCE: 21419/90368
: CURRENT APPLICATION NUMBER: US/09/844,705
: CURRENT FILING DATE: 2001-04-27
: PRIOR APPLICATION NUMBER: 09/443,766
: PRIOR FILING DATE: 1999-11-19
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 12
: LENGTH: 1369
: TYPE: DNA
: ORGANISM: Porcine
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (9)..(1103)
: US-09-844-705-12

```

	Query Match	36.9%;	Score 423.8;	DB 10;	Length 1269;
	Best Local Similarity	69.6%;	Pred. No. 2.8e+125;		
	Matches 591;	Conservative 0;	Mismatches 252;	Indels 6;	Gaps 1;
QY	217	GGCATGTTCACGATCAATTC	CATTGGCCGCTGGGAACCA	GCAGATGGCGGATACGCCACA	276
Db	243	GGACCTGGACIATTTACCG	GGATGGCGGTTTGGAAACCA	GATGGGACAGTATGCCACG	302
QY	277	CTCTTTTGCACTGGCCAGG	ATGAACGGACGCTTGCCTT	CATCTCCCGCATCCATCACAAC	336
Db	303	CTGCTGGCCCTGGCGAGCT	CAACGGCGCGACGCTTCAT	CTCCAGCTGCATCGACGCC	362
QY	337	GCCTCTAGCGCCCATCTTC	CAGGATCAGCCTCCCGGTG	TTACACGCGACGGGCCAANAAG	396
Db	363	GTCCCTGGCCCCCGTGTT	CCGCATCACGCTGCCCTG	CTCTGCGGCCGAGGTACACGCGAC	422
QY	397	ATCCGATGCGCAATTTACCA	CTCATACGACTGGANGAGG	GGGTTTACCGCCACATTCGG	456
Db	423	GTCCCTTGGCGGGAGTG	TGGAGCTTCACGACTGGATG	CTCCGAGATATGCGCCACTTAAG	482

457	QY	GGACATTTGTGGCTTCAGGGATACCCGTTGTCTTGACCTTTACCAACCACTCGGC	516
483	Db	GAGCCCTGGCTGAAGTCAACGGCTTCCGCTGCTCTGGACCTTCTCCACCACTCCGG	542
517	QY	CCAGAGATCCTGAAGCAGTTCACCTTCGATGACCAAGTGGCGGAGGAGGCCAGGCCTTC	575
543	Db	GAGCAGATCCGACGAGGTTCACCTGCGACGACCACTTCGGCAAGAGGCCCAAGGGGTA	602
577	QY	CTGCGT-----GGTCTGGGGGTGAATGGAGGACGCGGAGTAGTTTGTGGGTGTCCAT	630
603	Db	CTGAGTCAGTTCGCTACCCCGACAGGGGACCGCCCAAGCACTTCGTGGGGTCCAC	662
631	QY	GTGCGCCGAGGGGACATGTGATGTCTAGCCCTAATGTGTGGAAGCGGTGTGGGTGAC	690
663	Db	GTGCGCCGGGGACATCTGGGTGTGATGCCAAGCGCTGGAAGGGGTGTGTGGGTGAC	722
691	QY	CGGGGTACTCTGAAAGCCCTGGATATGTTCCGGGCAAGCTATTATCTCCAGACTTC	750
723	Db	GGCGGTACTCCAGCAGGCTATGGACTGGTTCGGGCCCGCATACGAAGCCCGCTTT	782
751	QY	GTGGTTACAAGCAACGGTATGGCTGGTCCGGGGAGAACATTAATGCTTCCCGAGAGAC	810
783	Db	GTGGTCACCAAGCAACGGCATGGAGTGGTCCGGAGAAACATGACACCTCCCGGGGGAC	842
811	QY	GTGGTCTTCGGGGCAATGATTTAGGGGTCCCAAGCCCAAGACATCGCGCTGCTCACC	870
843	Db	GTGATCTTTGCTGGCATGGCGGGAGGCGCGCCCGCAGGGACATTTGGCGTCTGGTG	902
871	QY	CAGTGCACCAACCAACATCATGACTATTGGGACCTTTGGGATTGGGCTGCTACCTGGCA	930
903	Db	CAGTGCACCAACCAACATCATGACCATTTGCACCTTCGGGTTCTGGGCGCGCTACCTTGGCT	962
931	QY	GGTGGTGATACCATCTACATTAGCCAATACAGCCTTCCGGATTTCTCGTTCTCTCAAGTC	990

```

Db      963  GGTGAGATACCACTACTACTTGGCTAACTTCACCCIGGCCACTTCACAGCTTCCTGAAGATC 1022
Qy      991  TTTHAAGCAGGAGGAGGAGCCTTCCTACCCGAAATGGGTGGGCATCCCTGGCGATCTGTCCCA 1050
Db      1023  TTTHAACCCGAGGCTGCTTCCTCCCGAGTGGTGGGCATTAATGCAGACTGTCTCCA 1082
Qy      1051  CTCCTTTAAG 1059
Db      1083  CTCGAGATG 1091

RESULT 12
US-10-963-9
; Sequence 9, Application US/10105963
; Publication No. US20030068818A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Denning, Chris
; APPLICANT: Clark, A. John
; TITLE OF INVENTION: Animal Tissue with Carbohydrate Antigens Compatible for Human
; TITLE OF INVENTION: Transplantation and a Carbohydrate Determinant Selection Syst
; TITLE OF INVENTION: Recombination
; FILE REFERENCE: 731/002
; CURRENT APPLICATION NUMBER: US/10/105,963
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: US 60/277,811
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 3373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (104)..(1201)
; OTHER INFORMATION:
; US-10-105-963-9

```

Query Match	35.2%;	Score 405;	DB 9;	Length 3373;	
Best Local Similarity	58.4%;	Pred. No. 4,3e-119;			
Matches 577;	Conservative	0;	Mismatches 260;	Indels 6;	Gaps 1;
QY	217	GGCATTGTCACGATCAATTCATTGGCCGGCTGGGACACGAGTGGCGAATACGCCACA	276		
Db	338	GGCACCTGGACTGTCACCCCAATGGCCGCTTGGTAATCAGATGGACAGTATGCCACG	397		
QY	277	CATCTTGCACCTGGCCAGGATGAACGACGGCTGCGTTTCATCCCGCATCCATCAGACAAC	336		
Db	398	CTGCTGGCTGTGCCCCAGCTCAACGGCCGCCGGCCCTTTATCTGCTCGTCGATGCATGCC	457		
QY	337	GTCTTAGCGCCCATCTTCAGGATCAGCCTCCCGGTGTTACACGCGACACGGCCAAAAAG	396		
Db	458	GCCTGCGCCCGGTATTCCGCATCACCTGCCCGTCTGCCCCCAGAGTGGACGCCGC	517		
QY	397	AUCCCATGGAGAAATACCATCTCACAGACTGATGGAGAGAGCGTTTACCGCCACATTCOG	456		
Db	518	AGCCCTGGCGGAGCTGCAGCTTCAGACTGATGTCCGAGGAGTACGGGACTTGAGA	577		
QY	457	GGACACTTTGTGGCTTCACGGGATACCGGTGCTCTGGACCTCTACACGACCTTCGCG	516		
Db	578	GATCCTTTCTGAAGCTCTGTGCTTCCCTGCTCTTGAGCTTTCTTCCACCATCTCCGG	637		
QY	517	CCAGAGATCCTGAAGGAGTTCACCTTCATGACCATGACGTCGGAGGAGGCCACGGCTTC	576		
Db	638	GAAACAGATCCGACAGAGTTACCTTCGACGACCACTTCGGGAAGAGGCCAGAGTGTG	697		
QY	577	CTGGGTGGTCTCGGGTG-----AATGGAGCCACCGGAGTACTTTTGTGGGTGTCCAT	630		
Db	698	CTGGGTGACTCGGCTTGGCGCGACAGGGGACCGCCCGCACTTTGTTCGGCGTCCAC	757		
QY	631	GTGCGCGGAGGAGACTATGTGATGTTCATGCCTTAATGTGTGAAGGCGGTGGTGGTGAC	690		

Db 758 GTGGCGGTGGGACTATCTCAGGTTATGCTTCAGGCGTGGAAAGGTTGGTGGGCGAC 817  
Qy 691 CGGGTTACCTGGAAAGGCGCTGGATATGTTCCGGGCGACGCTATCACTCCAGTCTTC 750  
Db 818 AGCGCTACCTCCGGCAGGCGCATGAGTGGTTCCGGGCGACGCGACGAAAGCGCGTTTC 877  
Qy 751 GTGGTTCAAGCAACGGTATGCGCTGGTGGCGGAGAACATTAATGCTTCCCGAGGAGAC 810  
Db 878 GTGTCACCAACAGCGCATGAGTGGTGTAAAGAAACATCGACACTCCAGGGCGAT 937  
Qy 811 GTGTTGTCGGGCGCATGTTATGAGGGTTCGGGCGACGCGCAAGGACTTGGCGTCTACTCACC 870  
Db 938 GTGAGGTTGCTGGCGATGGACAGAGGCTACACCGTGGAAAGACTTGGCCGTGTCACA 997  
Qy 871 CAGTGCACACACCATCATCATGACTATGGGACCTTGGGATTTGGCTTGGCTTGGCTTGGCT 930  
Db 998 CAGTGCACACACCATCATCATGACTATGGGACCTTGGGCTTGGCTTGGCTTGGCTTGGCT 1057  
Qy 931 GTGGTGATACCATCTACTTAGCCAACTACACCTTCCGGATTCTCCGTTTCCICAAAGTC 990  
Db 1058 GCGGAGACACTGCTACTCGGCAACTTACCTTCCGAGACTCTGAGTTCTTGAAGATC 1117  
Qy 991 TTTAAGCAGGAGGAGGCTTCTTACCGAATGGGTGGGCGATCCCTCGCGGATCTGTCCCCA 1050  
Db 1118 TTTAAGCGGAGGCGGCTTCTTCCGAGAGTGGGTGGGCGATTAATGAGACTTGTCTCCA 1177  
Qy 1051 CTC 1053  
Db 1178 CTC 1180

## RESULT 13

US-09-863-475A-5  
Sequence 5, Application US/09863475A  
Patent No. US20020102688A1  
GENERAL INFORMATION:  
APPLICANT: LOWE, JOHN B.  
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
GLYLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
P.C.  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/863,475A  
FILING DATE: 24-May-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/914,281  
FILING DATE: 20-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248955 OPAT JR  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:

LENGTH: 8174 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
ANTI-SENSE: NO  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-863-475A-5

Query Match 35.2%; Score 405; DB 10; Length 8174;  
Best Local Similarity 68.4%; Pred. No. 6.2e-119;  
Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;  
Qy 217 GGCACTGTTACAGATCAATTCGATTCGCGCGTGGGAAACAGATGGCGCATACCCACA 276  
Db 4920 GGCACCTGGAGTGTCTACCCCAATGGCGGTTTGGTAATCAGATGGGACAGTATGCCACG 4979  
Qy 277 CTCCTTTGACATGGCGAGGATGAACGAGCGGTTCGTTTCATCCCCCGCATCATGCACAA 336  
Db 4980 CTGCTGGCTCTGGCCAGCTCAACGGCGCGCGGCTTTATCTTCCCTGCCATGCATGCC 5039  
Qy 337 GCTTAGCGGCCATCTTCAGGATCAGCTCCCGGTGTTTACACAGGACACGGCCAAAAG 396  
Db 5040 GCCCTGGCGCGGATTCGGATCACCCTGCGCGTGTGGGCCACAGAAAGTGCACGCGC 5099  
Qy 397 ATCCCATGCGAGAAATTACCATCTCAACGACTGGATGGAGAGCGTTACCGCCACATCCG 456  
Db 5100 AGCGGTGGCGGAGCTGACGCTTACGACTGGATGCGGAGGATACCGGACTTTGAGA 5159  
Qy 457 GGACACTTGTGCGCTTACCGGGATACCCGTGCTCTCGAGCTTCTACACCACTCGGC 516  
Db 5160 GATCCTTCTCTGAAAGCTCTCTGGCTTCCCTGCTTGGACTTCTTCCACCATCTCGG 5219  
Qy 517 CCAGAGATCTGAAGGAGTTCACTCAGTGCATACACAGTCCGGGAGGAGGCGCCCTTC 576  
Db 5220 GAACAGATCCCGCAGAGAGTTCACTCAGTGCATACACCACTTCCGGGAGAGGCGCAGTGT 5279  
Qy 577 CTGGTGTGCTGCGGGTG-----AATGGAGCGACGCGAGTACTTTTGTGGTGTCCAT 630  
Db 5280 CTGGTGTGCTGCGGGTGCGGCGGCGGACGCGCGCGACCTTTGTGCGGCTCCAC 5339  
Qy 631 GTGCGCGGAGGCGACTATGTCATGCTATGCTTAAAGGCGGTGTGGCTGAC 690  
Db 5340 GTGCGCGTGGGAGTATCTCAGGTTATGCTCAGCGCTGAAAGGTTGTGGTGGCGAC 5399  
Qy 691 CGGGTTACCTGGAAGGCGCTGATGTTCCGGGACCGCTATCTATCTCCAGTCTTC 750  
Db 5400 AGCGCTTACCTCCGCGAGGCGTGGTTCGGGCGACGCGCGAGGCGCGCTTTC 5459  
Qy 751 GTGGTTACAAGCAACGGTATGGCTGTCGGGAGAACATTAATGCTTCCGAGGAGAC 810  
Db 5460 GTGTCACCAACAGCGCATGAGTGTGTAAAGAAACATCGACACTCCAGGGCGAT 5519  
Qy 811 GTGTTGTCGGGCAATGTTATGAGGGTTCGCCAGCGGAGGCTTGGCGTCTCACC 870  
Db 5520 GTGACGTTTGTGCGGATGGACAGAGGCTACACCGTGGAAAGACTTGGCCCTGCACACA 5579  
Qy 871 CAGTGCACACACCATCATGACTATTTGGGACCTTTGGGATTTGGGCTGCTACTCGCA 930  
Db 5580 CAGTGCACACACCATCATGACTATTTGGGACCTTGGGCTTCTGGGCTGCTACTCGCT 5639  
Qy 931 GGTGTGATACCATCTACTTAGCCAACTACACCTTCCGGATCTCCGTTCTCTCAAGTC 990  
Db 5640 GCGGAGACACTGCTTACCTGGGCAACTTACCTGCCAGCTCTGATTTCTCTGAAGATC 5699  
Qy 991 TTTAAGCAGGAGGCGCTTCTTACCCGAAATGGGTGGGCGATCCCTGCCGATCTGTCCCA 1050  
Db 5700 TTTAAGCGGAGGCGGCTTCTTCCCGAGTGGGTGGGCGATTAATGAGACTTGTCTCCA 5759  
Qy 1051 CTC 1053  
Db 5760 CTC 5762

RESULT 14  
US-09-999-672-29  
; Sequence 29, Application US/09999672  
; Patent No. US20020127653A1  
; GENERAL INFORMATION:  
; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: 8511-029  
; CURRENT APPLICATION NUMBER: US/09/999,672  
; CURRENT FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: US/09/298,886  
; PRIOR FILING DATE: 1999-04-26  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 29  
; LENGTH: 100  
; TYPE: DNA  
; ORGANISM: Oryctolagus cuniculus  
US-09-999-672-29

Query Match 6.3%; Score 72.8; DB 10; Length 100;  
Best Local Similarity 83.0%; Pred. No. 2.4e-13;  
Matches 83; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
QY 827 ATGGTATTGAGGGTCCGCGCAAGCAAGACTTCGGCTGCTCACCACCTGCAACCCACCA 886  
Db 1 ATGGCTTCGAGAGCTCGCGCGCAAGCAAGACTTCGGCTGCTCACCACCTGCAACCCACCG 60  
QY 887 TCATGACTATTGGGACCTTTGGGATTTGGGCTGCCTACCT 926  
Db 61 TCATGACCATCGGCACCTTTGGGATCTGGCGCCCTACCT 100

RESULT 15  
US-10-040-863-29  
; Sequence 29, Application US/10040863  
; Patent No. US20020137165A1  
; GENERAL INFORMATION:  
; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: 8511-029  
; CURRENT APPLICATION NUMBER: US/10/040,863  
; CURRENT FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: 09/298,886  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 29  
; LENGTH: 100  
; TYPE: DNA  
; ORGANISM: Oryctolagus cuniculus  
US-10-040-863-29

Query Match 6.3%; Score 72.8; DB 12; Length 100;  
Best Local Similarity 83.0%; Pred. No. 2.4e-13;  
Matches 83; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
QY 827 ATGGTATTGAGGGTCCGCGCAAGCAAGACTTCGGCTGCTCACCACCTGCAACCCACCA 886  
Db 1 ATGGCTTCGAGAGCTCGCGCGCAAGCAAGACTTCGGCTGCTCACCACCTGCAACCCACCG 60  
QY 887 TCATGACTATTGGGACCTTTGGGATTTGGGCTGCCTACCT 926  
Db 61 TCATGACCATCGGCACCTTTGGGATCTGGCGCCCTACCT 100

Search completed: May 26, 2003, 15:19:53  
Job time : 178.211 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 10:09:53 ; Search time 1886.5 Seconds  
(without alignments)  
9864.119 Million cell updates/sec

Title: US-10-040-863-7

Perfect score: 1149

Sequence: 1 atggcagcgccaggttcc.....gaagagccttcgatggaaa 1149

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Watch 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*\*

2: em\_esthum:\*\*

3: em\_estin:\*\*

4: em\_estnu:\*\*

5: em\_estov:\*\*

6: em\_estpl:\*\*

7: em\_estro:\*\*

8: em\_htc:\*\*

9: gb\_est1:\*\*

10: gb\_est2:\*\*

11: gb\_htc:\*\*

12: gb\_est3:\*\*

13: gb\_est4:\*\*

14: gb\_est5:\*\*

15: em\_estfun:\*\*

16: em\_estom:\*\*

17: gb\_gss:\*\*

18: em\_gss\_hum:\*\*

19: em\_gss\_inv:\*\*

20: em\_gss\_pln:\*\*

21: em\_gss\_vrt:\*\*

22: em\_gss\_fun:\*\*

23: em\_gss\_man:\*\*

24: em\_gss\_mus:\*\*

25: em\_gss\_other:\*\*

26: em\_gss\_pro:\*\*

27: em\_gss\_rod:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	410	35.7	656	10	BB624949
2	398.6	34.7	1133	14	BQ2020588
3	343.8	29.9	714	10	BE266792
4	283	24.6	596	14	BM744084
5	275.4	24.0	590	14	BM764661
6	273.2	23.8	591	14	BM781532

7	272.4	23.7	587	14	BM764788
8	271.8	23.7	784	9	AL046138
9	265.4	23.1	552	14	BM764751
10	247	21.5	521	12	BF710517
11	243	21.1	440	17	BB849545
12	232.2	20.2	449	17	BB274506
13	217	18.9	532	14	BM773356
14	212.2	18.5	895	10	BE299719
15	208.4	18.1	436	10	BB848859
16	207.8	18.1	541	12	BF194101
17	190.2	16.6	513	12	BE756417
18	187	16.3	326	10	AF843917
19	186.4	16.2	463	12	BF358052
20	176.8	15.4	838	12	BG192820
21	167.4	14.6	537	12	BE750891
22	159.6	13.9	207	10	AW347661
23	150.6	13.1	596	13	BJ091033
24	136.4	11.9	578	17	AZ285162
25	136	11.8	383	12	BF930329
26	122.4	10.7	424	13	BI534769
27	120.4	10.5	655	13	BJ040991
28	107.2	9.3	837	12	BF304028
29	106.8	9.3	521	17	BF342474
30	99.4	8.7	366	10	BE667589
31	96.8	8.4	526	10	BE019528
32	80	7.0	279	10	BB850127
33	75.8	6.6	649	10	AW917919
34	75.6	6.6	224	10	AW838837
35	75.6	6.6	484	13	BI350052
36	72.6	6.3	518	17	BF342476
37	67.6	5.9	264	14	BM742943
38	67.6	5.9	300	9	AU100082
39	67.6	5.9	304	9	AL709188
40	67.6	5.9	728	14	BM974044
41	65.6	5.7	554	12	BG610560
42	64.4	5.6	265	14	BM758349
43	57.4	5.0	541	12	BF418108
44	51.2	4.5	475	9	AI606683
45	50.4	4.4	578	9	AA108903

#### ALIGNMENTS

RESULT 1	BB624949	656 bp	mRNA	linear	EST 31-AUG-2001
LOCUS	BB624949	RIKEN full-length enriched, adult male colon	Mus musculus		
DEFINITION	CDNA clone 9030420111 5', mRNA sequence.				
ACCESSION	BB624949				
VERSION	BB624949.1	GI:15398432			
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
AUTHORS	1 (bases 1 to 656)				
	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takanashi,F., Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.				
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)				
JOURNAL	Unpublished (2001)				
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216				









ECORI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(drf)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F' with electroporation method."

```

BASE COUNT      123 a      201 c      153 g      113 t
ORIGIN
Query Match      24.0%; Score 275.4; DB 14; Length 590;
Best Local Similarity 71.2%; Pred. No. 3.1e-65;
Matches 394; Conservative 0; Mismatches 126; Indels 33; Gaps 1;

QY  9 CGCCAGGTTCCCTTCCTTCCTGCGCCACTTCCTCATCTTCTTCGTGACTTC 68
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  71 CGTTCAGATGCCCTTCCTTCCTTCCTGCGCCACTTCCTCATCTTCTTCGTG 130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  69 CACCATCATCCATCCATCCAGCAGGAATAGTGAAGCTCCCAACCCCTGCAGAGAAGGAATT 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  131 CACTATATTTTACGTTTCAGCAGCGGCTAGCGAAGATCAAGCCATGTGGAGTTACCGGT 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  129 ACCGATCAGCACTCAATGTCTTCGGGAACACAGAAAGCCAGAGATGCGACGGGACAG 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  191 GCAGATACCACTAGTCTACCCCTCAACATCAAAGGCACCTGGGACCCA----- 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  189 CGAGCAGCATGGGAATGGAGAGCTGCGGGGCATGTTCAAGATCAATTCATGGCGGCT 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  235 -----GCCAGCTCAGGGGATGTGAGCATCAATGCAATAGGCCGCT 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  249 GGGGAACAGATGGGGAATACGCCACACTTTTGCACTGGCCAGGATGAACGACGGCT 308
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  278 GGGGAACAGATGGGGAATACGCCACACTGTACGCCCTGGCCAAGATGAAGGGCGGCC 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  309 TGGTTTCATCCCGCATCCATGACACAGCTCTAGCCGCCATCTTCAGGATCAGCCTCC 368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  338 CGCCTTCATCCCGCCAGATGACACAGCCCTGGCCGCCATCTTCAGATCAGCCTGCC 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  369 GGTGTACACAGGCACACGGCCAAAAGATCCCATGGCAGAAATTAOCATCTCAACGACTG 428
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  396 GTGCTGCACAGGCCACGGCCAGGATCCCTGGCAGAACTACCACTGAATGACTG 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  429 GATGGAGAGCGTTACCGGCACATTCGGGACACTTGTGGCTTCACGGGATACCCGCTG 488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  458 GATGGAGAGGAATACCGGCACATTCGCGGGGAGTACGTCGCTTCACCGGCTACCCCTG 517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  489 CTCTGACCTTCTACCAACACCTGCGCCAGAGATCCTGAAGAGGTTTCACCCTGATGA 548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  518 CTCTGACCTTCTACCAACACCTGCGCCAGAGATCCTGACAGGAGTTTCACCCTGACGA 577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  549 CCAGTCCGGGAG 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  578 CCAGTCCGGGAG 590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 6
LOCUS   BM781532      591 bp      mRNA      linear      EST 05-MAR-2002
DEFINITION K-EST0058109 S3SN16s1 Homo sapiens cDNA clone S3SN16s1-19-B07 5',
            mRNA sequence.
ACCESSION BM781532

```

```

BM781532.1      GI:19129764
EST.
SOURCE     Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 591)
            Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            K.M.Y.S.
            21c Frontier Korean EST Project 2001
            Unpublished (2002)
            Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 19 row: B column: 07
            High quality sequence stop: 591.
            Location/Qualifiers
                1..591
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="S3SN16s1-19-B07"
                /clone_lib="S3SN16s1"
                /sex="F"
                /tissue_type="Ascites"
                /cell_type="Lymphoblast-like"
                /lab_host="SNU-16"
                /lab_host="Top10F'"
                /note="Organ: Stomach; Vector: pTZ18Rpl; Site_1: EcoRI;
                Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                bacterial alkaline phosphatase (BAP) and then decapped
                with tabacco acid pyrophosphatase (TAP). The decapped
                intact mRNA was ligated with DNA-RNA linker including EcoR
                I site by treatment of T4 RNA ligase and the first strand
                cDNA was synthesized from oligo dt-selected mRNA by
                priming with dt-tailed vector. The dt-tailed vector was
                adjusted to have about 60nt. The cDNA vector was
                circularized with E. coli DNA ligase after digestion of
                EcoRI which site is also included in vector. An RNA strand
                converted to a DNA strand by Okayama-Berg method. The
                obtained cDNA vectors were used for transformation of
                competent cells E. coli Top10F' by electroporation method.
                The cDNA libraries constructed by this method are
                full-length enriched cDNA library. After analyzing and
                sequencing about 2,000 ~ 3,000 colonies in original cDNA
                library, the abundant cDNAs were selected and amplified by
                PCR reaction using vector region primer including T7
                promoter as 5' primer and N(drf)14 as 3' primer. The PCR
                products were used as template for synthesis of
                biotinylated single stranded RNA by in vitro transcription
                reaction. The synthesized RNA probes were hybridized with
                antisense single stranded cDNAs prepared from original
                library and incubated with avidin-gel. After removing
                DNA-RNA hybrids by centrifuge, the subtracted cDNA
                libraries were constructed by transformation of the
                remaining DNA into competent cells E. coli Top10F' with
                electroporation method."
BASE COUNT      123 a      201 c      150 g      117 t
ORIGIN
Query Match      23.8%; Score 273.2; DB 14; Length 591;
Best Local Similarity 70.9%; Pred. No. 1.3e-64;
Matches 393; Conservative 0; Mismatches 128; Indels 33; Gaps 1;

QY  9 CGCCAGGTTCCCTTCCTTCCTGCGCCACTTCCTCATCTTCTTCGTGACTTC 68
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  71 CGTTCAGATGCCCTTCCTTCCTTCCTGCGCCACTTCCTCATCTTCTTCGTGACTTC 130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  69 CACCATCATCCATCCAGCAGGAATAGTGAAGCTCCCAACCCCTGCAGAGAAGGAATT 128

```



```

QY      549 CCACGTGGCG 558
Db      578 CCACGTGGCG 587

AL046138      784 bp mRNA linear EST 29-FEB-2000
DKFP434P0672_r1 434 (synonym: htes3) Homo sapiens cDNA clone
DKFP434P0672 5', mRNA sequence.
AL046138
AL046138.2 GI:5936190
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 784)
Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
EST (Poustka, et al.)
Unpublished (1999)
On Jul 9, 1999 this sequence version replaced gi:5434225.
Contact: Poustka A.J.
Department Lehrach
Max-Planck-Institute for Molecular Genetics
Innestrasse 73, 14195 Berlin, Germany
Tel: +49-30-84131623
Fax: +49-30-84131128
Email: poustka@mpg-berlin-dahlem.mpg.de
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Resequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFP434P0672) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES             source
Location/Qualifiers
1..784
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFP434P0672"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="PH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      142 a 234 c 248 g 157 t
ORIGIN
Query Match      23.7%; Score 271.8; DB 9; Length 784;
Best Local Similarity 73.6%; Pred. No. 3.6e-64;
Matches 360; Conservative 0; Mismatches 127; Indels 2; Gaps 1;

QY      555 GCGGAGGAGGCCAGCCCTCTCGCTGGTGTGGGGTGAATGGAGCCAGCCGAGTAC 614
Db      1 GCGCGAGGAGGCCAGAACTTCTCGGGGCCCTGCGAGGCCAAGTGGGAGCGGACGGC 60

QY      615 TTITGTGGTGCCATGTGCGCGGAGGAGCATGTGCATGTCCCTAAATGTGTGGAA 674
Db      1 GCGCGAGGAGGCCAGAACTTCTCGGGGCCCTGCGAGGCCAAGTGGGAGCGGACGGC 60

QY      61 CTCTGTGGGGTCCAGCTGCGCGGGGGGACATGTCCGTGTGATGTGCGCGCTATGGAA 120
Db      1 GCGCGAGGAGGCCAGAACTTCTCGGGGCCCTGCGAGGCCAAGTGGGAGCGGACGGC 60

QY      675 GGGCGTGTGGTGGACCGGGGTACCTGGAAGAAGGCCCTGGATATGTTCCGGGACGGCA 734
Db      1 GCGCGAGGAGGCCAGAACTTCTCGGGGCCCTGCGAGGCCAAGTGGGAGCGGACGGC 60

QY      121 GGGGGTGTCTGGCGAGCC--GGTACCTGACGGGGGCGCTGGAGTGTTCGGGCGCTGCTG 178
Db      1 GCGCGAGGAGGCCAGAACTTCTCGGGGCCCTGCGAGGCCAAGTGGGAGCGGACGGC 60

QY      735 TTCATCTCCAGTCTTCTGCTGTACACAGCAACGGTATGGCTGTGCGCGGAGCAACATTAA 794
Db      1 GCGCGAGGAGGCCAGAACTTCTCGGGGCCCTGCGAGGCCAAGTGGGAGCGGACGGC 60

QY      179 CGCCCTCCGGGCTTTGTGTGTACACAGCATGACATGGGCTGTGTCGGGAGAGCATCAA 838
Db      1 GCGCGAGGAGGCCAGAACTTCTCGGGGCCCTGCGAGGCCAAGTGGGAGCGGACGGC 60

QY      795 TGCITCCCGAGGAGAGCTGTGTCTCGGGGCAATGTAITGAGGGGTGCGCCACCAAGGA 854
Db      1 GCGCGAGGAGGCCAGAACTTCTCGGGGCCCTGCGAGGCCAAGTGGGAGCGGACGGC 60

```



```
sex=mixed), (tissue_type=hippocampus, dev_stage=adult,
sex=male), (tissue_type=whole body, dev_stage=9 days
embryo, sex=mixed), (tissue_type=lung, dev_stage=13 days
embryo, sex=mixed)"
103 a 134 c 99 g 104 t
BASE COUNT
```

Query Match	21.1%;	Score 243;	DB 10;	Length 440;
Best Local Similarity	83.5%;	Pred. No. 2.5e-56;		
Matches 308;	Conservative	0;	Mismatches 40;	Indels 21;
				Gaps 2

61	GTGACTTCCACATATCCACCTCCAGCAGGGAATAGTGAAGCTCCAAACCCCTGTGCAGAG	120
Db	GTGACTTCCACATATCCACCTCCAGCAGGGAATAGTGAAGCTCCAAACCCCTGTGCAGAG	207
Qy	121 AAGGAATACCGATGACGACTCAATGTCCTCGGGAACACACAAAGCCACAGATGGCA	180
Db	208 AAGGAATATAC--AGGCGGTTCATATGCTCTACCAACCGCGCAAGACAGACAT----	260

QY 301 GGACGGGTTGCGTTATCCCGCGATCCATGCACACGGTCTAGGCCCATCTTCAGGATC 360

db 367 GGTGGGTTGGCTTCATCCCTGAATCCATGCACACGGTCTAGGCCCATCTTCAGGATC 426

RESULT 12  
BH274506  
LOCUS  
DEFINITION  
449 bp DNA linear GSS 30-NOV-2000  
CH330-23L10.TJ CHORT-230 Segment 1 Rattus norvegicus genomic clone  
CH330-23L10, DNA sequence.  
ACCESSION  
BH274506  
VERSION  
BH274506.1 GI:17186908  
KEYWORDS  
GSS.

Other\_GSSs: CH230-23L10.TV  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel.: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the rat BAC library CHORI-230  
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library  
availability, please contact Pieter de Jong ([pdejong@mail.cho.org](mailto:pdejong@mail.cho.org))  
Clones may be purchased from BACPAC Resources  
(<http://www.chori.org/bacpac/orering/information.htm>). BAC end  
page: [http://www.tigr.org/tdb/bac\\_ends/rat/bac\\_end\\_tiro.html](http://www.tigr.org/tdb/bac_ends/rat/bac_end_tiro.html)

Location/Qualifiers  
1. .449

```

ity sequence sup:0322.
Location/Qualifiers
1..532
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S3SNU16s1-17-H05"
/clone_lib="S3SNU16s1"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast"
/cell_line="SNU-16"
/lab_host="Wron16"

```

/tissue\_type="Ascites"  
 /cell\_type="Lymphoblast-like"  
 /cell\_line="SNU-16"  
 /lab\_host="Top10F"  
 /note="organ: Stomach; Vector: pTZ18RP1; Site\_1: EcoRI;  
 Site\_2: NotI; The poly (A)+ RNA was dephosphorylated with  
 bacterial alkaline phosphatase (BAP) and then decapped  
 with tabacco acid pyrophosphatase (TAP). The decapped  
 intact mRNA was ligated with DNA-RNA linker including EcoRI  
 I site by treatment of T4 RNA ligase and the first strand  
 cDNA was synthesized from oligo dt-selected mRNA by  
 priming with dt-tailed vector. The dt-tailed vector was  
 adjusted to have about 60nt. The cDNA vector was  
 circularized with E. coli DNA ligase after digestion of  
 EcoRI which site is also included in vector. An RNA strand  
 converted to a DNA strand by Okayama-Berg method. The  
 obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10F' by electroporation method.  
 The cDNA libraries constructed by this method are  
 full-length enriched cDNA library. After analyzing and  
 sequencing about 2,000 ~ 3,000 colonies in original cDNA  
 library, the abundant cDNAs were selected and amplified by  
 PCR reaction using vector region primer including T7  
 promoter as 5' primer and N(dt)14 as 3' primer. The PCR  
 products were used as template for synthesis of  
 biotinylated single stranded RNA by in vitro transcription  
 reaction. The synthesized RNA probes were hybridized with  
 antisense single stranded cDNAs prepared from original  
 library and incubated with avidin-gel. After removing  
 DNA-RNA hybrids by centrifuge, the subtracted cDNA  
 libraries were constructed by transformation of the  
 remaining DNA into competent cells E. coli Top10F' with  
 electroporation method."

Site 2: NotI; The poly (+) RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dr-selected mRNA by priming with dr-tailed vector. The dr-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 - 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and NdrT14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F' with electroporation method."

174 c 1136 g 103 t

BASE COUNT	109 a	174 c	136 g	103 t
ORIGIN	electrophoretation method.			
Query Match	18.9%; Score 217; DB 14; Length 522;			
Best Local Similarity	68.5%; Pred. No. 4.4e-49;			
Matches 332; Conservative	0; Mismatches 120; Indels 33; Gaps 1;			

Matches	332	Conservative	0	Mismatches	120	Indels	33	Gaps	17
---------	-----	--------------	---	------------	-----	--------	----	------	----

  

Qy	9	CGGCCAGTTCCTTCCTTCCTCGGCCACTTCCTCATCTTTGTCTTCGACTTC	68
Db	71	CGTTCAGATGCCCTTCCTTCCTCCCATCGGCCACTTCATCCTCTTTCTCTTACGGTTTC	130

QY 69 CACCATCATCCACTCCAGCAGCGAATAGTAGCTCCAACCCCTGTCACAGAGGAATT 138  
||| || | | | | | | | | | | | | | | | | | |  
Db 131 CACTATATTTCAAGTTCACGACGGGTAGCGAGATTCAGGCATGTGGGAGTTACGGT 190

Qy	129	ACCGATGACGACTCAATGTCTCGGAAACACAGAAAGCCAGAGATGCGGGGACAG	188
Db	191	GCAGATACCAAGTCTAGCCTCAACATCAAAAGGCATGGGACCA	234
Qy	189	CGASCACATGGGAATGGGATGGGGGCATGTCACGATCAATTCATTGGCGGCT	248
	235	CGAGATGACGACTCAATGTCTCGGAAACACAGAAAGCCAGAGATGCGGGGACAG	277

DB	235	-----GCGAGCTCAGGGGGGATGTGACCGATCAATGCAATAGCGCGCCT	277
QY	249	GGGGACCAGATGGCGCAATAGCCGACACTTTTGGACTGGCCAGGATGACCGGACGCGCT	308

Db	278	GGGGAACAGATGGGGAGTACGCCACACTGTACGCCCTGGCCAGATGAACGGGGCGCC	337
DV	309	TGGGTTCATCCCGCCATCCATTCACACAGCGCTAGCGCCCATCTCTAGGAACAGCGTCCC	368

```

338 CGCCTTCATCCCGCCGACAGATGACAGCACCCCTGGCCCCCATCTTCAGAAATCAACCTGCC 397
369 GTGTTACACAGGACAGCCGCAAAAGATCCCATGCGCAGAAATACCATCTCAACGACTG 428
398 GTGCTGCACAGGCGCCACGCGCAGAGATCCCTTGGCAGAACTACCACTGAATGACTG 457
429 GATGAGGAGCGTTACCGCCACATCTCCGGGACACTTTGTGCGCTTACGGGATACCGGTG 488
458 GATGAGGAGGAAATACCGCACTTCCGGGGAGTAGCTCGCTTACCGGCTACCCGCTAC 517
489 CTCCT 493
518 CTCCT 522

RESULT 14
BE299719
LOCUS
DEFINITION BE299719 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2960318 5',
mRNA sequence.
ACCESSION BE299719
VERSION BE299719.1 GI:9183467
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 895)
AUTHORS NIH-MGC http://mgc.nhl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLC53 row: h column: 15
High quality sequence stop: 758.
FEATURES
Location/Qualifiers
1..895
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2960318"
/clone_lib="NIH_MGC_17"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: muscle; Vector: pOTB7; Site:1: EcoRI;
Site:2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

331 CACAACGCTCTAGCGCCCATCTTCAGGATCAGCTCCCGGTGTTTACAGGACACGGCC 390
482 CACAGCACCTGGCCGCCCATCTTCAGATCACCCTGCCGTGTCACAGCGCCACGGCC 541
391 AAAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGAGCGCTTACCGCAC 450
542 AGCAGGATCCCTGGCAGAACTACCACTGAACGACTGGATGGAGGAGGAGTACCGGCAC 601
451 ATTCCGGGACACT-TTGTGCGCTTACGGGAT-ACCGTGTCTCTCTGACCT---TCFACC 505
602 ATCCCGGGGCGCTGANGTCCGCTCACGGCTAACCCCTGCTCTCTGACGACTCTATACCA 661
506 ACCACCTGCGCCGACAGATCC-TGAAGGAGTTTACCCTGC-ATGACACAGTGGGGAGGA 563
662 CCAACCTCCGCGAGGAGATCTTCCAGGAGTTTACCCTGCAACGACACCGTGGCGGAGGA 721
564 GGCCGAGCGCTTCTCGGTGGTCTGCGGGTGAATGGGAGCCAGCCGAGTACTTTTGGG 623
722 GGCCAGAAAGTTCTCGGGGCGCTGCGAGGCCAAGTGGCGAGGCGGAGGACGACTTCTGG 781
624 TGTCATGTGCGCGGAGGAGGACTATGTGCACTG 656
782 GTTCC-CGTCCGCGGGGAGGACTATGTCCGGGT 813

RESULT 15
BB848859
LOCUS
DEFINITION BB848859 RIKEN full-length enriched, adult inner ear Mus musculus
cDNA clone F930004014 5', mRNA sequence.
ACCESSION BB848859
VERSION BB848859.1 GI:17090313
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 436)
AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.,
Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Iweda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
```





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:52:27 ; Search time 48.1951 Seconds  
(without alignments)  
1050.630 Million cell updates/sec

Title: US-10-040-863-8

Perfect score: 2032

Sequence: 1 MAAQVFFSPLAHFLIFVF.....RSHFLKAKGVTCYVAGRAF 380

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

1:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2032	100.0	380	21 AAB36104	Rat hepatoma H35 C
2	2032	100.0	380	23 AAE16622	Rat hepatoma H35 C
3	1896	93.3	353	21 AAB36105	Rat hepatoma H35 C
4	1896	93.3	353	23 AAE16623	Rat hepatoma H35 C
5	1502	73.9	340	19 AAW37855	Porcine secretor t
6	1502	73.9	340	19 AAW53101	Pig secretor. Sus
7	1485.5	73.1	343	19 AAW69332	Human Sec2 protein
8	1475	72.6	344	21 AAB36106	Human Sec2 catalyt
9	1057.5	52.0	365	12 AAB36104	GDP-Fuc:beta-D-gal
10	1057.5	52.0	365	16 AAR80154	GDP-L-fucose-beta-

11	1057.5	52.0	365	16 AAR70422	2-Alpha-fucosyltra
12	1057.5	52.0	365	17 AAR70421	2-Alpha-fucosyltra
13	1057.5	52.0	365	17 AAR90572	Human H-transferrin
14	1057.5	52.0	385	18 AAR23805	Human alpha(1,2) fu
15	1057.5	52.0	385	18 AAW13640	Human alpha(1,2)-f
16	1057.5	52.0	365	21 AAY97279	Human H-transferrin
17	1049.5	51.6	365	15 AAR45936	A glycosyltransfer
18	1046.5	51.5	365	19 AAW53102	Pig H transferase.
19	1046.5	51.5	365	21 AAY79302	Pig alpha-1-2 fucc
20	1046.5	51.5	385	23 AAB47995	Swine alpha(1,2)
21	1044.5	51.4	385	20 AAW97356	Swine alpha(1,2) f
22	1044.5	51.4	385	20 AAW30630	Swine alpha(1,2) f
23	995	49.0	357	20 AAY17969	Swine alpha-1,2-fu
24	328.5	16.2	110	21 AAG03530	X. laevis alpha-1.
25	141	6.9	300	21 AAY92713	Human secreted pro
26	141	6.9	300	22 AAG64871	H. pylori UA802 al
27	139.5	6.9	287	23 AAW51992	Helicobacter pylor
28	103	5.1	661	22 AAG39722	Bacteroides fragil
29	101	5.0	690	23 AAE20142	Propionibacterium
30	97	4.8	317	13 AAR26061	Human protein cont
31	96.5	4.7	355	21 AAY70408	Growth Factor Rece
32	95	4.7	1515	22 AAG10824	C. elegans alpha-1
33	93.5	4.6	402	22 AAB63790	Novel human diagno
34	93	4.6	1073	22 AAG28599	Drosophila melanog
35	92	4.5	984	22 AAG93762	Novel human diagno
36	90.5	4.5	659	22 AAB46426	Human polypeptide.
37	90.5	4.5	2008	23 AAU11814	Billin binding prot
38	90.5	4.5	2306	23 AAU11817	Cancer and neuroge
39	89.5	4.5	2352	23 AAU11816	Cancer and neuroge
40	89.5	4.4	586	21 AAY54126	Amino acid sequenc
41	89.5	4.4	604	22 AAU46700	Propionibacterium
42	89	4.4	353	16 AAB61993	Porphyromonas geng
43	89	4.4	334	22 AAB93020	Human protein sequ
44	89	4.4	1031	22 AAU17082	Novel signal trans
45	89	4.4	2057	22 AAE11890	Angiogenesis assoc

#### ALIGNMENTS

RESULT 1  
AAB36104  
ID AAB36104 standard; Protein; 380 AA.  
XX  
AC AAB36104;  
XX  
DT 19-FEB-2001 (first entry)  
XX  
DE Rat hepatoma H35 cell alphas-2-fucosyltransferase.  
XX  
DE Rat; alphas-2-fucosyltransferase; cytotstatic; neuroprotective;  
KW nootropic; gene therapy; Fucalpal-2Galbeta1-3GalNAc; immunotherapy;  
KW immunosuppression; cancer; neurological disease;  
KW Small cell lung carcinoma.  
XX  
OS Rattus norvegicus.  
XX  
PN WO200064464-A1.  
XX  
PD 02-NOV-2000.  
XX  
PF 23-APR-1999; 99WO-US07384.  
XX  
PR 23-APR-1999; 99WO-US07384.  
PA (PACI-) PACIFIC NORTHWEST CANCER FOUND.  
XX  
PI Holmes EH, Sherwood AL;  
XX WPI: 2000-687262/57.  
DR N-PSDB; AAC67965.  
XX  
PT New rat ganglioside GM1-specific alphas-2-fucosyltransferase, useful for

PT Preparation of fucosyl GM1 which is useful as a nutritional composition  
 XX or immunotherapeutic for cancer and neurological diseases -  
 PS Claim 1; Fig 5; 91pp; English.

CC The present sequence was given in a specification relating to an isolated  
 CC rat ganglioside GM1-specific alpha1-2fucosyltransferase protein. The  
 CC protein or its cellular fraction is useful for synthesis of a molecule  
 CC comprising Fucalpal-2Galbeta1-3GalNAc, a glycolipid, glycoprotein,  
 CC glycolipoprotein or a free oligosaccharide comprising  
 CC Fucalpal-2Galbeta1-3GalNAc. The method involves contacting  
 CC alpha1-2fucosyltransferase with GDP-fucose and a molecule or glycolipid,  
 CC glycoprotein, glycolipoprotein or oligosaccharide having a terminal  
 CC Galbeta1-3GalNAc group. It is also useful for synthesis of fucosyl-GM1  
 CC by contacting the protein with GDP-fucose and ganglioside GM1. The  
 CC obtained glycoproteins, glycolipoproteins, glycolipids and  
 CC oligosaccharides are useful as nutritional compositions and fucosyl-GM1  
 CC is useful for inducing an immunotherapeutic or immunosuppressive action  
 CC against cancer, neurological disease or small cell lung carcinoma.

XX Sequence 380 AA;

Query Match 100.0%; Score 2032; DB 21; Length 380;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-212;  
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAQVFFSPFLAHFLIFVFTSTIIHQORIVKLOPLSEKELPMTQMSSGNTSPENR 60

DB 1 MASAQVFFSPFLAHFLIFVFTSTIIHQORIVKLOPLSEKELPMTQMSSGNTSPENR 60

QY 61 RSEQHGNGELRGMTINSIGRLGNMGGEYATLFALARMNGRLAFTPASMHNALAPIFRI 120

DB 61 RSEQHGNGELRGMTINSIGRLGNMGGEYATLFALARMNGRLAFTPASMHNALAPIFRI 120

QY 121 SLPLVLSHTAKKIPWQNYHLNDWMEERYRHIPGHVFRFTGPCSWTFYHHLRPEILKEFT 180

DB 121 SLPLVLSHTAKKIPWQNYHLNDWMEERYRHIPGHVFRFTGPCSWTFYHHLRPEILKEFT 180

QY 181 LHDHVREAAQAFRLGRVNGSQPSTFVGHVHVRGDIYVHVMNVKGVADRGYLEKALDM 240

DB 181 LHDHVREAAQAFRLGRVNGSQPSTFVGHVHVRGDIYVHVMNVKGVADRGYLEKALDM 240

QY 241 FRARYSSPVFTVTSNGMAWCRENINASRGDVVFAGNIEGSPAKDFALLTQCNHTIMTG 300

DB 241 FRARYSSPVFTVTSNGMAWCRENINASRGDVVFAGNIEGSPAKDFALLTQCNHTIMTG 300

QY 301 TFGIWAAYLAGDITIYLANYTLPDSPFLKVFKEPAAFLEPWVGIPADLSPLLKALTPACP 360

DB 301 TFGIWAAYLAGDITIYLANYTLPDSPFLKVFKEPAAFLEPWVGIPADLSPLLKALTPACP 360

QY 361 RSHFHLKAGVTCYVAGRAF 380

DB 361 RSHFHLKAGVTCYVAGRAF 380

RESULT 2

ID AAE16622 standard; Protein; 380 AA.

XX AAE16622;

XX AAE16622;

DT 09-APR-2002 (first entry)

XX Rat hepatoma H35 cell alpha1-2Fuct.

DE Rat;

XX Rat;

XX Rat;

KW galactose beta1-3N-acetylgalactosamine; Galbeta1-3GalNAc; glycolipid;

KW glycoprotein; glycolipoprotein; oligosaccharide; fucosyl-GM1; cancer;

KW gene therapy; oncogenic transformation; cytostatic; ganglioside; GM1;

XX cell transformation.

OS Rattus norvegicus.

XX

FH Key Location/Qualifiers  
 FT Domain 1..27  
 FT /note= "Intracellular/Transmembrane domain"  
 FT Modified-site 199  
 FT /note= "N-glycosylated"  
 FT Region 221..380  
 FT /note= "Region which overlaps rat FTB"  
 FT Modified-site 265  
 FT /note= "N-glycosylated"  
 FT Modified-site 293  
 FT /note= "N-glycosylated"  
 XX US6329170-B1.  
 PN 11-DEC-2001.  
 XX 23-APR-1999; 99US-0298886.  
 XX 23-APR-1999; 99US-0298886.  
 XX (WHO-) NORTHWEST HOSPITAL.  
 XX Holmes EH, Sherwood AL;  
 XX WPI: 2002-121132/16.  
 XX N-PSDB; AAD27207.  
 XX Rat hepatoma H35 cell alpha1-2fucosyltransferase, useful for producing  
 GM1-specific alpha1-2fucosyltransferase enzyme by recombinant  
 techniques and for detecting oncogenic transformation of test tissues -  
 Claim 3; Fig 5; 41pp; English.  
 XX The invention relates to rat GM1-specific alpha1-2fucosyltransferase  
 (alpha1-2Fuct) enzyme and its corresponding nucleic acid. This nucleic  
 acid is specific for a carbohydrate moiety found in ganglioside GM1,  
 a terminal galactose beta1-3N-acetylgalactosamine (Galbeta1-3GalNAc);  
 saccharide. Alpha1-2Fuct DNA is useful for producing rat alpha1-2Fuct  
 protein by recombinant techniques. Alpha1-2Fuct DNA is useful for the  
 preparative synthesis of fucosyl containing glycolipids, glycoproteins,  
 glycolipoproteins and oligosaccharide, and for preparing fucosyl-GM1.  
 Alpha1-2Fuct DNA is useful for detecting oncogenic transformation which  
 involves assaying for changes in expression of alpha1-2 Fuct. Since  
 alpha1-2Fuct is activated in cell transformation, antisense sequences  
 derived from alpha1-2Fuct DNA are useful for inhibiting, suppressing  
 or treating cancer. Alpha1-2Fuct DNA is useful in gene therapy and  
 antisense therapy. The present sequence is rat hepatoma H35 cell  
 alpha1-2Fuct.  
 XX Sequence 380 AA;  
 Query Match 100.0%; Score 2032; DB 23; Length 380;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-212;  
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASAQVFFSPFLAHFLIFVFTSTIIHQORIVKLOPLSEKELPMTQMSSGNTSPENR 60  
 DB 1 MASAQVFFSPFLAHFLIFVFTSTIIHQORIVKLOPLSEKELPMTQMSSGNTSPENR 60  
 QY 61 RSEQHGNGELRGMTINSIGRLGNMGGEYATLFALARMNGRLAFTPASMHNALAPIFRI 120  
 DB 61 RSEQHGNGELRGMTINSIGRLGNMGGEYATLFALARMNGRLAFTPASMHNALAPIFRI 120  
 QY 121 SLPLVLSHTAKKIPWQNYHLNDWMEERYRHIPGHVFRFTGPCSWTFYHHLRPEILKEFT 180  
 DB 121 SLPLVLSHTAKKIPWQNYHLNDWMEERYRHIPGHVFRFTGPCSWTFYHHLRPEILKEFT 180  
 QY 181 LHDHVREAAQAFRLGRVNGSQPSTFVGHVHVRGDIYVHVMNVKGVADRGYLEKALDM 240  
 DB 181 LHDHVREAAQAFRLGRVNGSQPSTFVGHVHVRGDIYVHVMNVKGVADRGYLEKALDM 240  
 QY 241 FRARYSSPVFTVTSNGMAWCRENINASRGDVVFAGNIEGSPAKDFALLTQCNHTIMTG 300  
 DB 241 FRARYSSPVFTVTSNGMAWCRENINASRGDVVFAGNIEGSPAKDFALLTQCNHTIMTG 300  
 QY 301 TFGIWAAYLAGDITIYLANYTLPDSPFLKVFKEPAAFLEPWVGIPADLSPLLKALTPACP 360  
 DB 301 TFGIWAAYLAGDITIYLANYTLPDSPFLKVFKEPAAFLEPWVGIPADLSPLLKALTPACP 360  
 QY 361 RSHFHLKAGVTCYVAGRAF 380  
 DB 361 RSHFHLKAGVTCYVAGRAF 380

Db 241 FRARYSPVFTVTSNGMWCRENINASRGDVVFAGNGIEGSPAKDFALLTQCCHNTIMTIG 300  
QY 301 TFGIWAAYLAGGDTIYLANITLPDPSFLKVFKEPAFLPEWVGIPADLSPLLKALTPACP 360  
Db 301 TFGIWAAYLAGGDTIYLANITLPDPSFLKVFKEPAFLPEWVGIPADLSPLLKALTPACP 360  
QY 361 RSHFHLKAGVTCYVAGRAF 380  
Db 361 RSHFHLKAGVTCYVAGRAF 380

RESULT 3  
AAB36105  
ID AAB36105 standard; Protein; 353 AA.  
XX AC  
XX AAB36105;  
DT 19-FEB-2001 (first entry)  
XX DE  
XX Rat hepatoma H35 cell alphas-2fucosyltransferase catalytic domain.  
KW Rat; alphas-2fucosyltransferase; cytostatic; neuroprotective;  
KW nontropic; gene therapy; Fucalphas-2galactosamine; Galbetal-3galNAC; immunotherapy;  
KW immunosuppression; cancer; neurological disease;  
KW small cell lung carcinoma.  
XX  
OS Rattus norvegicus.  
XX  
XX WO200064464-A1.  
PD 02-NOV-2000.  
XX  
XX 23-APR-1999; 99WO-US07384.  
XX  
XX 23-APR-1999; 99WO-US07384.  
XX  
XX (PACI-) PACIFIC NORTHWEST CANCER FOUND.  
XX  
XX Holmes EH, Sherwood AL;  
XX  
XX WPI; 2000-687262/67.  
XX N-PSDB; AAC67966.  
XX  
XX New rat ganglioside GM1-specific alphas-2fucosyltransferase, useful for  
XX preparation of fucosyl GM1 which is useful as a nutritional composition  
XX or immunotherapeutic for cancer and neurological diseases -  
XX  
XX Claim 2; Fig 3A; 9lpp; English.  
XX  
XX The present sequence is given in a specification relating to a rat  
XX ganglioside GM1-specific alphas-2fucosyltransferase protein. The protein  
XX or its cellular fraction is useful for synthesis of a molecule comprising  
XX Fucalphas-2galbetal-3galNAC, a glycolipid, glycoprotein, glycolipoprotein  
XX or a free oligosaccharide comprising Fucalphas-2galbetal-3galNAC.  
XX The method involves contacting alphas-2fucosyltransferase with GDP-fucose  
XX and a molecule or glycolipid, glycoprotein, glycolipoprotein or  
XX oligosaccharide having a terminal Galbetal-3galNAC group. It is also  
XX useful for synthesis of fucosyl-GM1 by contacting the protein with  
XX GDP-fucose and ganglioside GM1. The obtained glycoproteins,  
XX glycolipoproteins, glycolipids and oligosaccharides are useful as  
XX nutritional compositions and fucosyl-GM1 is useful for inducing an  
XX immunotherapeutic or immunosuppressive action against cancer,  
XX neurological disease or small cell lung carcinoma.  
XX  
XX Sequence 353 AA;  
XX

Query Match 93.3%; Score 1896; DB 21; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1e-197;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2B LQQRIVKLOPLSEKELPMTOMSSGNTSPSPMRDSEHGNGELRGMTINSIGRLGNQM 87  
Db 1 LQQRIVKLOPLSEKELPMTOMSSGNTSPSPMRDSEHGNGELRGMTINSIGRLGNQM 60

QY 88 GEYATLFAALARMNGRLAIFASMHENALAPIFRISLPLVHSDTAKKIPWQNYHLNDQMEER 147  
Db 61 GEYATLFAALARMNGRLAIFASMHENALAPIFRISLPLVHSDTAKKIPWQNYHLNDQMEER 120  
QY 148 YRHIPGHVRFETGYPGCSWTIFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSIFV 207  
Db 121 YRHIPGHVRFETGYPGCSWTIFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSIFV 180  
QY 208 GYHVRRGDXYHVMNPVWKGVVADRGYLEKALDMFRARYSSPVFVTSNGMWCRENINAS 267  
Db 181 GYHVRRGDXYHVMNPVWKGVVADRGYLEKALDMFRARYSSPVFVTSNGMWCRENINAS 240  
QY 268 RGDVVFAGNGIEGSPAKDFALLTQCCHNTIMTIGTGGIWAAYLAGGDTIYLANITLPDPSF 327  
Db 241 RGDVVFAGNGIEGSPAKDFALLTQCCHNTIMTIGTGGIWAAYLAGGDTIYLANITLPDPSF 300  
QY 328 LKVFKEPAFLPEWVGIPADLSPLLKALTPACPRSHFHLKAGVTCYVAGRAF 380  
Db 301 LKVFKEPAFLPEWVGIPADLSPLLKALTPACPRSHFHLKAGVTCYVAGRAF 353

RESULT 4  
AAE16623  
ID AAE16623 standard; Protein; 353 AA.  
XX AC  
XX AAE16623;  
DT 09-APR-2002 (first entry)  
XX  
XX Rat hepatoma H35 cell alphas-2fucosyltransferase catalytic domain.  
KW Rat; alphas-2fucosyltransferase; alphas-2fucosyltransferase; alphas-2fucosyltransferase;  
KW galactose betal-3-N-acetylglucosamine; Galbetal-3galNAC; glycolipid;  
KW glycoprotein; glycolipoprotein; oligosaccharide; fucosyl-GM1; cancer;  
KW gene therapy; oncogenic transformation; cytostatic; ganglioside; GM1;  
KW cell transformation; catalytic domain.  
XX  
XX Rattus norvegicus.  
XX  
XX  
XX Key Location/Qualifiers  
XX Modified-site 172  
XX Region 194..353  
XX Modified-site 238  
XX Modified-site 286  
XX Modified-site 292  
XX Modified-site 292  
XX US6329170-B1.  
XX 11-DEC-2001.  
XX 23-APR-1999; 99US-0298886.  
XX 23-APR-1999; 99US-0298886.  
XX (NWHO-) NORTHWEST HOSPITAL.  
XX Holmes EH, Sherwood AL;  
XX WPI; 2002-121132/16.  
XX N-PSDB; AAD27208.  
XX  
XX Rat hepatoma H35 cell alphas-2fucosyltransferase, useful for producing  
XX GM1-specific alphas-2fucosyltransferase enzyme by recombinant  
XX techniques and for detecting oncogenic transformation of test tissues -  
XX  
XX Claim 4; Fig 3; 4lpp; English.  
XX

CC The invention relates to rat GM1-specific alpha1-2fucosyltransferase  
 CC (alpha1-2fucT) enzyme and its corresponding nucleic acid. This nucleic  
 CC acid is specific for a carbohydrate moiety found in ganglioside GM1,  
 CC a terminal galactose beta1-3N-acetylgalactosamine (Galbeta1-3GalNAc),  
 CC saccharide. Alpha1-2fucT DNA is useful for producing rat alpha1-2fucT  
 CC protein by recombinant techniques. Alpha1-2fucT DNA is useful for the  
 CC preparative synthesis of fucosyl containing glycolipids, glycoproteins,  
 CC glycolipoproteins and oligosaccharide, and for preparing fucosyl-GM1.  
 CC Alpha1-2fucT DNA is useful for detecting oncogenic transformation which  
 CC involves assaying for changes in expression of alpha1-2 fucT. Since  
 CC alpha1-2fucT is activated in cell transformation, antisense sequences  
 CC derived from alpha1-2fucT DNA are useful for inhibiting, suppressing  
 CC or treating cancer. Alpha1-2fucT DNA is useful in gene therapy and  
 CC antisense therapy. The present sequence is rat hepatoma H35 cell  
 CC alpha1-2fucT catalytic domain.

XX Sequence 353 AA;

Query Match 93.3%; Score 1896; DB 23; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 1e-197;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LQORIVKLPSEKELPMTQSSGNTSPENRRDSEHGNGELRGMTINSIGRLGNOM 87  
 DB 1 LQORIVKLPSEKELPMTQSSGNTSPENRRDSEHGNGELRGMTINSIGRLGNOM 60

QY 88 GYATLFLALRMNGRLAIPASMHNALAIPRISLPLVHSDTAKKIPQNYHLNDWMEER 147  
 DB 61 GYATLFLALRMNGRLAIPASMHNALAIPRISLPLVHSDTAKKIPQNYHLNDWMEER 120

QY 148 YRHIPGHEVRTGTPCSTFVHHLRPETLKEFTLHDHYREAZAQLRLGRVNGSPSTFV 207  
 DB 121 YRHIPGHEVRTGTPCSTFVHHLRPETLKEFTLHDHYREAZAQLRLGRVNGSPSTFV 180

QY 208 GVHVRGSDYVHMPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCENINAS 267  
 DB 181 GVHVRGSDYVHMPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCENINAS 240

QY 268 RGDVVFAAGNGTEGSPAKDFALLTCQNHITIMTIGFIWAAYLAGSDTTYLANYLTPDSPF 327  
 DB 241 RGDVVFAAGNGTEGSPAKDFALLTCQNHITIMTIGFIWAAYLAGSDTTYLANYLTPDSPF 300

QY 328 LKVKPEAAFLPEWGVIPADLSPILKALTACPRSHFLKAKGVTCYVAGRAF 380  
 DB 301 LKVKPEAAFLPEWGVIPADLSPILKALTACPRSHFLKAKGVTCYVAGRAF 353

RESULT 5  
 AAW37855

DB AAW37855 standard; Protein; 340 AA.

XX AC AAW37855;

XX DT 28-AUG-1998 (first entry)

XX XX Porcine secretor transferase (FUT2).

XX KW Secretor; glycosyltransferase; FUT2; pig; epitope; antigen;  
 XX KW transgenic animal; xenotransplantation; organ transplant.

XX OS Sus scrofa.

XX FT Key Location/Qualifiers

XX FT Domain 1..4 /note= "N-terminal cytoplasmic tail"

XX FT Domain 5..26 /note= "transmembrane domain"

XX FT Domain 27..340 /note= "C-terminal domain"

XX FT Modified-site 185..187 /note= "Asn is N-glycosylated"

XX FT Modified-site 251..253 /note= "Asn is N-glycosylated"

XX FT Modified-site 353 /note= "Asn is N-glycosylated"

FT Modified-site 279..281 /note= "Asn is N-glycosylated"

FT Modified-site 305..307 /note= "Asn is N-glycosylated"

XX WO9807837-Al.

XX PD 26-FEB-1998.

XX PF 22-AUG-1997; 97WO-AU00540.

XX PR 23-AUG-1996; 96AU-0001823.

XX PA (AUST-) AUSTIN RES INST.

XX PI McKenzie IFC, Sandrin MS;

XX DR WPI; 1998-169148/15.

XX DR N-PSDB; AAV29003.

XX PT Nucleic acid encoding glycosyltransferase able to compete with  
 XX PT second such enzyme - particularly used to reduce expression of  
 XX PT unwanted carbohydrate epitope(s) on tissues intended for  
 XX PT transplantation

XX PS Claim 6; Fig 1A-B; 40pp; English.

XX CC This polypeptide comprises porcine secretor glycosyltransferase  
 CC (SE or FUT2), a type II integral membrane protein has high affinity  
 CC for type I and type III substrates. Its amino acid sequence was  
 CC deduced from the nucleotide sequence of a genomic DNA clone (see  
 CC AAV29003) isolated from a pig liver library on the basis of homology  
 CC to human FUT2 cDNA. Pig FUT2 shows 83.2% identity with human FUT2.  
 CC Expression of a glycosyltransferase, such as porcine Se, in a  
 CC tissue results in reduced expression of unwanted carbohydrate  
 CC epitopes on the tissue, especially porcine heart, liver, kidney or  
 CC pancreas, rendering it more suitable for transplantation, i.e. less  
 CC immunogenic and of increased immunological acceptability. A  
 CC claimed method of producing a cell from a donor species that is  
 CC immunologically acceptable to a recipient species involves reducing  
 CC levels of carbohydrate on the donor cell that causes it to be  
 CC recognised as non-self by the recipient by expressing a nucleic  
 CC acid for a glycosyltransferase such as porcine Se in the cell.

XX XX Sequence 340 AA;

Query Match 73.9%; Score 1502; DB 19; Length 340;  
 Best Local Similarity 79.6%; Pred. No. 9.1e-155;  
 Matches 281; Conservative 28; Mismatches 30; Indels 14; Gaps 3;

QY 1 MASAQVPSEFFLAHLFIQVVTSTIIHLQQRIVKLQPLSEKELPMTQSSGNTSPENR 60  
 DB 1 MLSMQASPFPTGPFILFVFTASTIFHLQQRVMTQP--TWELQWVQT--TSP--- 52

QY 61 RDSQHGNGELRGMTINSIGRLGNOMGEYATLALARMNGRLAIPASMHNALAIPRI 120  
 DB 53 -----SSPQLKGMWTINAIGRLGNOMGEYATLALARMNGRPAFIPPEMHTLAPRI 106

QY 121 SLPLVHSDTAKKIPQNYHLNDWMEERYRHIPGHFVRTGTPCSTFVHHLRPETLKEFT 180  
 DB 107 TLPVLHASTARRIPQNYHLNDWMEERYRHIPGVRVTGTPCSTFVHHLRPETLKEFT 166

QY 181 LHDHVREREAQFLRGLRVNGSQPSTFVGHVVRGSDYVHMPNVKGVVADRGYLEKALDM 240  
 DB 167 LHNHVREREAQDFLRGLRVNGSRPSTYVGVHVRGSDYVHMPNVKGVVADRGYLEKALDW 226

QY 241 FRARYSPFVVTNSGMWCENINASRGDVFVAGNGIEGSPAKDFALLTCQNHITIMTIG 300  
 DB 227 FRARYSPFVVTNSGMWCENINASRGDVFVAGNGIEGSPAKDFALLTCQNHITIMTIG 286

QY 301 TFGIWAAYLAGSDTTYLANYLTPDSPFLKVKPEAAFLPEWGVIPADLSPILK 353  
 DB 287 TFGIWAAYLAGSDTTYLANYLTPDSPFLKVKPEAAFLPEWGVIPADLSPILK 339

```

RESULT 6
AAW53101
ID AAW53101 standard; Protein; 340 AA.
XX
AC AAW53101;
XX
DT 08-JUL-1998 (first entry)
XX
DE Pig secretor.
XX
KW Pig; secretor; chimeric; glycosyltransferase; gene therapy;
transplantation.
XX
OS Sus scrofa.
XX
PN WO9805768-A1.
XX
PD 12-FEB-1998.
XX
PF 01-AUG-1997; 97WO-AU00492.
XX
PR 21-AUG-1996; 96US-0024279.
XX
PR 02-AUG-1996; 96AU-0001402.
XX
PA (AUST-) AUSTIN RES INST.
XX
PI McKenzie IFC, Sandrin MS;
XX
DR WPI; 1998-159170/14.
DR N-PSDB; AAV21639.
XX
Nucleic acids encoding chimeric glycosyltransferases - used for
altering carbohydrate levels on the surface of cells, useful in gene
therapy and transplantation
XX
Example 2; Fig 6; 51pp; English.
XX
The present sequence represents pig secretor used in an example of the
present invention. The present invention describes nucleic acids (NA)
encoding a chimeric glycosyltransferase. The NAs comprise a catalytic
domain of a first glycosyltransferase (GT) and a localisation signal of
a second GT, whereby when the NA is expressed in a cell and where the
chimeric enzyme is located in an area of the cell where it is able to
compete for substrate with a second GT, resulting in reduced levels of
a product from the second GT. The NAs can be used to produce cells and
organs with desired glycosylation patterns. Products and methods of the
present invention can be used to reduce the levels of undesirable
epitopes in cells, tissues or organs which may be used in
transplantation or gene therapy.
XX
Sequence 340 AA;
XX
Query Match 73.9%; Score 1502; DB 19; Length 340;
Best Local Similarity 79.6%; Pred. No. 9.1e-155;
Matches 281; Conservative 28; Mismatches 30; Indels 14; Gaps 3;
QY 1 MASAQVFFFPFLAHFLIFVFTVTIIHLQQRIVKLOPLSEKELPMTQMSSGNTESPMR 60
DB 1 MLSMQASFFFGPFLFTFTASTIFHLQQRVKTQP--TWELQMTQVT---TESP--- 52
QY 61 RDSQHGNGELRGMTFINSIGRLGNOMGEYATLALARMNGRLAFIPASMHNALAPIFRI 120
DB 53 -----SSPQLKGMWTINAGRLGNOMGEYATLYALARMNGRPAFIPPMHSTLAPIFRI 106
QY 121 SLFVLHSDTAKKIPQWNYHLNDWMEERYHIFGCHYVRFTGFCPSWTFYHHLRPEILKEPT 180
DB 107 TLPVLHASTARRIPQWNYHLNDWMEERYHIFGCVYRLGFCPSWTFYHHLRTEILREPT 166
QY 181 LHDHVEEAQAFLRLRVNGSQPSTFVGHVHRGDYVHYMPNVNWKGVADRYLEKALDM 240
DB 167 LHNHVEEAQAQFLRLRVNGSRPSTYVGHVHRGDYVHYMPNVNWKGVADRYLEQALDW 226
QY 241 FRARYSSPVFVVTSGNAWCRNINASRGDVVFAGNGIEGSPAKDFALLTQCNTHTMTIG 300
DB 227 FRARYRSPVFVVTSSNGMAWCRNINASRGDVVFAGNGIEGSPAKDFALLTQCNTHTMTIG 286
QY 301 TFGIWAAYLAGGDTIYLANVTLTLPDSPFLKVFKEPAFAFLPEWVGIPADLSPLLK 353
DB 287 TFGIWAAYLAGGETIYLANVTLTLPDSPFLKLFKPEAAFLPEWIGIEADLSPLLK 339
RESULT 7
AAW69332
ID AAW69332 standard; Protein; 343 AA.
XX
AC AAW69332;
XX
DT 20-NOV-1998 (first entry)
XX
DE Human Sec2 protein sequence.
XX
KW Sec2; alpha(1,2) fucosyltransferase; H blood group; secretor genotyping;
GDP-L-fucose;beta-D-galactoside 2-alpha-L-fucosyltransferase; human;
KW FUT2; nonsecretor genotyping.
XX
OS Homo sapiens.
XX
PN US5807732-A.
XX
PD 15-SEP-1998.
XX
PF 28-FEB-1995; 95US-0395800.
XX
PR 28-FEB-1995; 95US-0395800.
XX
(GIOR/) GIORGI D.
PA (KELL/) KELLY R J.
PA (LENN/) LENNON G.
PA (LOWE/) LOWE J B.
PA (ROUQ/) ROUQUET S.
XX
Giorgi D, Kelly RJ, Lennon G, Lowe JB, Rouquier S;
WPI; 1998-520127/44.
DR N-PSDB; AAV58323.
XX
DNA encoding fucosyltransferase enzyme - useful for producing
recombinant enzyme and genotyping person as secretor or nonsecretor
XX
Disclosure; Column 45-50; 55pp; English.
XX
This sequence is the human Sec2 protein of the invention. The DNA
encodes a alpha(1,2) fucosyltransferase and is the Secretor
CC alpha(1,2)fucosyltransferase locus, that cross hybridises with the
CC H blood group alpha(1,2)fucosyltransferase gene. The DNA is useful for
CC producing a recombinant human GDP-L-fucose:beta-D-galactoside
CC 2-alpha-L-fucosyltransferase (FUT2) which can be used for genotyping an
CC individual as a secretor or nonsecretor as it is known that nonsecretors
CC homozygous for a mutant allele of the FUT2 gene that has a stop codon in
CC the position corresponding to amino acid 143.
XX
Sequence 343 AA;
XX
Query Match 73.1%; Score 1485.5; DB 19; Length 343;
Best Local Similarity 77.9%; Pred. No. 5.8e-153;
Matches 275; Conservative 30; Mismatches 37; Indels 11; Gaps 2;
QY 1 MASAQVFFFPFLAHFLIFVFTVTIIHLQQRIVKLOPLSEKELPMTQMSSGNTESPMR 60
DB 1 MLVQMPFSEPNHFLIFVFTVTSTIFHVQORLAKIQAM--WELPV-----QIPVLA 49
QY 61 RDSQHGNGELRGMTFINSIGRLGNOMGEYATLALARMNGRLAFIPASMHNALAPIFRI 120
DB 50 SYSKALGPSQLRGMTWTINAGRLGNOMGEYATLYALAKMNGRPAFIPQMHSTLAPIFRI 109

```





```

XX PI Cummings RD, Kopchik JJ, Moremen KW, Mukerji P;
XX PR Pierce JM, Prieto PA, Smith DF;
XX XX WPI; 1995-328284/42.
XX DR N-PSD3; AAT01083.
XX XX New transgenic non-human mammal milk prods - contg. heterologous
XX PT components produced as secondary gene prods. of an heterologous gene
XX PT Example 1; Page 62-64; 83pp; English.
XX PS
XX CC 2-Alpha-fucosyltransferase (AAR70421) is encoded by a cDNA clone
XX CC (AAT01083) isolated from a human epidermal carcinoma A431 library.
XX CC The enzyme can be expressed in the milk of a transgenic mammal,
XX CC esp. cow. This allows large-scale prodn. of oligosaccharides and
XX CC glycosylated proteins and lipids in the milk.
XX XX
XX SQ Sequence 365 AA;
XX
XX Query Match 52.0%; Score 1057.5; DB 16; Length 365;
XX Best Local Similarity 56.5%; Pred. No. 3.1e-106;
XX Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps 4;
XX
QY 12 LAHLIFVFTSTIIHLQ-----RIVKLOPLSEKELPMTQMSSGNTE 56
DB 11 LAFILVCLSVIFFLHIDQSPFHGLGLSLCPDRRLVTPPPVAIFCLPGTAMGNASSC 70
QY 57 PEMRSEQHGNGELGMFTINSIGRLGNOMGEYATLALARMNGRLAFTPASHNALAP 116
DB 71 P-----QH-PASLSGTWVYPNGRFGNQMGQYATLLAQLNGRAFTLPAMHAALAP 122
QY 117 IFRISLPVLSDTAKKIPWNYHLNDWMEERYRHIPGHFVFTGYPCSWTFYHHLRPEIL 176
DB 123 VFRITLPVLAPEVDSRTPRELQHDWMSEYADLRDPFLKLSGFCSTFFHHLRQIR 182
QY 177 KETLHDHVREREAQAFRLGRV--NGSQPSTFVGHVHRRGDYVHVMPNVKGVVADRGYL 234
DB 183 REFTLHDHVREREAQVGLQRLGTGDRPRTFVGHVHRRGDYLVQVMPQKRVGVSAYL 242
QY 235 EKALDMFRARYSSPVFVTSNGMAWCRENINASKGDVVFAGNGIEGSPAKDFALLTQCNE 294
DB 243 RQAMDWFRARHEAPVFTVTSNGMECKENIDTSQGDVTFAGDQGEATPWKDFALLTQCNE 302
QY 295 TIMTIGTGFHWAAYLAGGDTIYLYANTLPDSPFLKVPKPEAAFLPEWVGIPADLSPLKA 354
DB 303 TIMTIGTGFHWAAYLAGGDTVLYANFTLPDSEFLKIFKPEAAFLPEWVGINADLSPLWL 362
QY 355 LTP 357
DB 363 AKP 365
XX
RESULT 12
AAR70421
ID AAR70421 standard; Protein; 365 AA.
XX AC AAR70421;
XX XX
XX DT 27-FEB-1996 (first entry)
XX DE 2-Alpha-fucosyltransferase.
XX KW GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyltransferase;
XX KW glycosyltransferase; oligosaccharide; glycoprotein; glycolipid;
XX KW transgenic animal; cattle; glycosylation; milk.
XX OS Homo sapiens.
XX XX
XX PN WO9524488-A1.
XX PD 14-SEP-1995.
XX XX

```

```

PF 24-JAN-1995; 95WO-US01147.
XX XX
PR 09-MAR-1994; 94US-0209132.
XX XX
PA (ABBO ) ABBOTT LAB.
XX XX
PI Cummings RD, Kopchik JJ, Moremen KW, Mukerji P;
PI Pierce JM, Prieto PA, Smith DF;
XX XX
DR WPI; 1995-328279/42.
DR N-PSD5; AAT01082.
XX XX
XX Transgenic animal expressing heterologous catalyst - used in
XX PT metabolite prodn esp. glycosyltransferase for prodn. of
XX PT oligosaccharide(s) and glyco-conjugate(s) in the milk, useful in
XX PT foods, pharmaceuticals, etc.
XX XX
XX Example 1; Page 62-64; 84pp; English.
XX XX
XX Alpha-1,2-fucosyltransferase (AAR70421) is encoded by a cDNA clone
XX CC (AAT01082) isolated from a human epidermal carcinoma cell line A431
XX CC cDNA library. The enzyme can be expressed in the milk of a
XX CC transgenic mammal, esp. cow. This allows large-scale prodn. of
XX CC oligosaccharides or glycosylated proteins and lipids in the milk.
XX XX
XX SQ Sequence 365 AA;
XX
XX Query Match 52.0%; Score 1057.5; DB 16; Length 365;
XX Best Local Similarity 56.5%; Pred. No. 3.1e-106;
XX Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps 4;
XX
QY 12 LAHLIFVFTSTIIHLQ-----RIVKLOPLSEKELPMTQMSSGNTE 56
DB 11 LAFILVCLSVIFFLHIDQSPFHGLGLSLCPDRRLVTPPPVAIFCLPGTAMGNASSC 70
QY 57 PEMRSEQHGNGELGMFTINSIGRLGNOMGEYATLALARMNGRLAFTPASHNALAP 116
DB 71 P-----QH-PASLSGTWVYPNGRFGNQMGQYATLLAQLNGRAFTLPAMHAALAP 122
QY 117 IFRISLPVLSDTAKKIPWNYHLNDWMEERYRHIPGHFVFTGYPCSWTFYHHLRPEIL 176
DB 123 VFRITLPVLAPEVDSRTPRELQHDWMSEYADLRDPFLKLSGFCSTFFHHLRQIR 182
QY 177 KETLHDHVREREAQAFRLGRV--NGSQPSTFVGHVHRRGDYVHVMPNVKGVVADRGYL 234
DB 183 REFTLHDHVREREAQVGLQRLGTGDRPRTFVGHVHRRGDYLVQVMPQKRVGVSAYL 242
QY 235 EKALDMFRARYSSPVFVTSNGMAWCRENINASKGDVVFAGNGIEGSPAKDFALLTQCNE 294
DB 243 RQAMDWFRARHEAPVFTVTSNGMECKENIDTSQGDVTFAGDQGEATPWKDFALLTQCNE 302
QY 295 TIMTIGTGFHWAAYLAGGDTIYLYANTLPDSPFLKVPKPEAAFLPEWVGIPADLSPLKA 354
DB 303 TIMTIGTGFHWAAYLAGGDTVLYANFTLPDSEFLKIFKPEAAFLPEWVGINADLSPLWL 362
QY 355 LTP 357
DB 363 AKP 365
XX
RESULT 13
AAR90572
ID AAR90572 standard; Protein; 365 AA.
XX AC AAR90572;
XX XX
XX DT 08-APR-1996 (first entry)
XX DE Human H-transferase.
XX KW H-transferase; xenograft hyperacute rejection; transplantation;
XX KW glycosyltransferase; galactose alpha(1,3) galactose.
XX XX

```



```

OS Homo sapiens.
XX
PN W09534202-A1.
XX
XX 21-DEC-1995.
XX
XX 14-JUN-1995; 95WO-US07554.
XX
XX 21-JUL-1994; 94US-0278282.
XX
XX 15-JUN-1994; 94US-0260201.
XX
XX (ALEX-) ALEXION PHARM INC.
XX
XX (AUST-) AUSTIN RES INST.
XX
XX
XX Fodor WL, Mckenzie IFC, Rother RP, Sandrin MS, Squinto SP;
XX
XX WPI; 1996-049326/05.
XX
XX N-PSDB; AAT12238.
XX
XX Redn. of rejection of xenogeneic cells following transplantation
XX by introducing a vector expressing fucosyltransferase into the
XX cells
XX
XX Example 1; Page 45-47; 69pp; English.
XX
XX The human H-transferase (AAR90572) product of a cDNA clone (AAT12238)
XX prepd. from human epidermoid carcinoma cells (ATCC CRL 155 A-431)
XX can be expressed in xenogeneic organs, tissues and cells using
XX a vector such as pAPEX-1 (AAT12239). This results in decreased
XX expression of the non-human antigen galactose alpha(1,3) galactose
XX on the surface of the organs etc. so that hyperacute rejection is
XX reduced upon transplantation to humans.
XX
XX Sequence 365 AA;
XX
XX Query Match 52.0%; Score 1057.5; DB 17; Length 365;
XX Best Local Similarity 56.5%; Pred. No. 3.1e-106;
XX Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps 4;
XX
QY 12 LAHFLIFVFTSTIIHQ-----RIVKLQPLSEKELPMTQMSSGNTE 56
DB 11 LAFLVCVLSVIFFLIHQDSFPHGLGLSILCPDRVTPPVAFICLPGTAGMGNASSC 70
QY 57 PEMRDSQHGNGELRGMTINSTIGRLGNQMGEXATLAFALARMNGRLAFIPASMHNALAP 116
DB 71 P-----QH-PASLSGTWTVYPNGRFGNMGQVATLLAQLAGRRAFILPAMHAALAP 122
QY 117 IFRISLPVLHSDTAKKIPWQNYHLNDWMEERYRHPGHFVFTGTPCSTWTFYHHLRPEIL 176
DB 123 VFRITLPVLAPEVDSRTPEWRELQLDHDMSEYADLRDPFLKLSGFPCCSWTFPHHLRQIR 182
QY 177 KEFTLHDHVREEAQAFRLGLRV--NGSQPSTFGVHVRRGDYVHVMPNVKGVVADRGYL 234
DB 183 REFTLHDHREEAQSVLQQLRGTRGDRPRTFVGVHVRRGDYLVMPQVWGVVGDSEAYL 242
QY 235 EKALDMFRARYSSPVVVTNSGMWACRENINASRGDVVFAGNGIEGSPAKDFALLTCQNH 294
DB 243 RQAMDWFRARHAPVFTVNSGMEWCKENIDTSQDVTFFAGDQGEATPWKDFALLTCQNH 302
QY 295 TIMTIGTGGWAAVLAGDGTIYLANFTLPDSFELKVKPEAAFLPEWVGIPADLSPLIKA 354
DB 303 TIMTIGTGGWAAVLAGDGTIYLANFTLPDSFELKVKPEAAFLPEWVGIPADLSPLIKA 354
QY 355 LTP 357
DB 363 AKP 365
XX
XX RESULT 14
XX AAW23805
XX ID AAW23805 standard; Protein; 365 AA.
XX
XX AC AAW23805;

```

```

XX
XX 15-SEP-1997 (first entry)
XX
XX Human alpha 1,2 fucosyltransferase (EC 2.4.1.69).
XX
XX Human GDP-L-fucose:beta-D-galactoside-2-alpha-L-fucosyltransferase;
XX alpha 1,2 FT; transgene; transgenic mouse; animal model;
XX intestinal adhesion; Helicobacter pylori infection; stomach;
XX small intestine; gut; epithelial cell; surface receptor;
XX carbohydrate antigen; gastritis; peptic ulcer; neoplasia;
XX gastric adenocarcinoma; Lewis antigen; fucosylation.
XX
XX Homo sapiens.
XX
XX US5625124-A.
XX
XX 29-APR-1997.
XX
XX 11-JUL-1994; 94US-0273411.
XX
XX 11-JUL-1994; 94US-0273411.
XX
XX (UNIW ) UNIV WASHINGTON.
XX
XX Falk P, Gordon JI;
XX
XX WPI; 1997-258275/23.
XX
XX N-PSDB; AAT178768.
XX
XX Animal model for Helicobacter pylori infection - comprising
XX transgenic mouse expressing human enzyme promoting intestinal
XX adhesion
XX
XX Example 1; Columns 13-16; 24pp; English.
XX
XX A claimed transgenic mouse expresses, in its intestinal epithelial
XX cells, the enzyme human GDP-L-fucose:beta-D-galactoside 2-alpha-L-
XX fucosyltransferase (also called alpha 1,2 FT) or human GDP-L-fucose:
XX beta-D-N-acetylglucosamide 3/4-alpha-L-fucosyltransferase (also
XX called alpha 1,3/4 FT). The enzyme is expressed under the
XX control of a gut epithelial cell-specific promoter and Helicobacter
XX pylori adheres to the transgenic cells. The transgenic mouse and
XX intestinal epithelial cells from it are useful as models for screening
XX compounds for the ability to inhibit adhesion of H. pylori to gut
XX epithelial cells. The present sequence represents human alpha 1,2 FT
XX and was published by Larsen et al., Proc.Natl.Acad.Sci.USA, 87:
XX 6674-6678 (1990).
XX
XX Sequence 365 AA;
XX
XX Query Match 52.0%; Score 1057.5; DB 18; Length 365;
XX Best Local Similarity 56.5%; Pred. No. 3.1e-106;
XX Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps 4;
XX
QY 12 LAHFLIFVFTSTIIHQ-----RIVKLQPLSEKELPMTQMSSGNTE 56
DB 11 LAFLVCVLSVIFFLIHQDSFPHGLGLSILCPDRVTPPVAFICLPGTAGMGNASSC 70
QY 57 PEMRDSQHGNGELRGMTINSTIGRLGNQMGEXATLAFALARMNGRLAFIPASMHNALAP 116
DB 71 P-----QH-PASLSGTWTVYPNGRFGNMGQVATLLAQLAGRRAFILPAMHAALAP 122
QY 117 IFRISLPVLHSDTAKKIPWQNYHLNDWMEERYRHPGHFVFTGTPCSTWTFYHHLRPEIL 176
DB 123 VFRITLPVLAPEVDSRTPEWRELQLDHDMSEYADLRDPFLKLSGFPCCSWTFPHHLRQIR 182
QY 177 KEFTLHDHVREEAQAFRLGLRV--NGSQPSTFGVHVRRGDYVHVMPNVKGVVADRGYL 234
DB 183 REFTLHDHREEAQSVLQQLRGTRGDRPRTFVGVHVRRGDYLVMPQVWGVVGDSEAYL 242
QY 235 EKALDMFRARYSSPVVVTNSGMWACRENINASRGDVVFAGNGIEGSPAKDFALLTCQNH 294
DB 243 RQAMDWFRARHAPVFTVNSGMEWCKENIDTSQDVTFFAGDQGEATPWKDFALLTCQNH 302

```

```

QY 295 TIMTIGTGGWAAYLAGGDTIYLANVTLPDSPLKVKFKEAAFLPEWVGTPADLSPLLKA 354
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 303 TIMTIGTGGWAAYLAGGDTIYLANVTLPDSPLKVKFKEAAFLPEWVGINADLSPLWTL 362
QY 355 LTP 357
Db 363 AKP 365

RESULT 15
AAW13640
ID AAW13640 standard; Protein; 365 AA.
XX AC
XX AAW13640;
XX 19-JUN-1997 (first entry)
XX Human alpha(1,2)-fucosyltransferase.
XX Alpha(1,2)-fucosyltransferase; glycosylation; oligosaccharide;
KW blood group H.
XX Homo sapiens.
XX WO9709421-A1.
XX 13-MAR-1997.
XX 06-SEP-1996; 96WO-US13816.
XX 08-SEP-1995; 95US-0525058.
XX (UNMI ) UNIV MICHIGAN.
XX Legault DJ, Lowe JB;
XX WPI; 1997-192897/17.
XX N-PSDB; AAT61677.
XX New recombinant fucosyltransferase proteins - useful for modifying
XX cell surface oligosaccharide structures
XX Example 1; Page 280-281; 329pp; English.
XX Human GDP-Fuc:beta-D-galactoside alpha(1,2)-fucosyltransferase
XX (AAW13640) is an enzyme involved in the expression of type I and II
XX blood group H structures. Its amino acid sequence was deduced
XX from an isolated DNA (AAT61677) derived from human A431 cells.
XX Expression of the fucosyltransferase animal cell lines provides
XX specific capabilities with respect to post-translational
XX modification of the oligosaccharides of expressed proteins or
XX lipids. Specific applns. of the enzyme include enzymatic
XX fucosylation of chain-terminating galactose residues on
XX lactosamine or neolacto type beta-D-galactoside to alpha-2-
XX fucose residues. The enzyme can also be used to raise antibodies
XX as diagnostic reagents and to screen cpds. for fucosyltransferase
XX inhibitor activity.
XX Sequence 365 AA:

```

```

Query Match 52.0%; Score 1057.5; DB 13; Length 365;
Best Local Similarity 56.5%; Pred. No. 3.1e-106;
Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps 4;

```

```

QY 12 LAHFLIFVFWTSTIIHQ-----RIVKLPLSEKELPMTTOMSSCNTES 56
||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 11 LAFLVCLSVIFFHIIHODSTPHGLGLSILCPDRRLVTPPVAIFCLPTAMGNASSC 70
57 PEMRDSGEQNGELGEMFTINSIGRLGNQMGYATIFALARNGRLAIPASMHNALAP 116
||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 71 P-----QH-PASLSGTWTVYPNGRFNGMGQYATILALAQNGERRAIFLPAHAAALAP 122

```

```

QY 117 IFRISLPVLHSOTAKKIPWNYHLNDWMERYRHIPGHFVFRFTGYPCSWTFYHHLRPEIL 176
:||||| : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 123 VFRTITPLVLAPEVDSPTPRRELQLDHWMSEYADLRDPFLKLSGFFCSTWTFHHLREQIR 182
QY 177 KEFTLHDHVREEAQAFLRGLRV--NGSQPSTFVGVHVRGDYVHVMPNWKGVYVADRGYL 234
:|||||:|||||: ||| : |||:|||||:|||||: ||| ||||| |||
Db 183 REFTLHDHLREEAQSVLGQLRLGRTGDRPRTFVGVHVRGDYLVQVMPQRWKGVVGSAYL 242
QY 235 EKALDMFRARYSSPVEFVTSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTQCNIH 294
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 243 RQAMDFRARHEAPVFEVTSNGMEWCKENIDTSQGDVTFAGDQGEATPWKDFALLTQCNIH 302
QY 295 TIMTIGTGGWAAYLAGGDTIYLANVTLPDSPLKVKFKEAAFLPEWVGTPADLSPLLKA 354
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 303 TIMTIGTGGWAAYLAGGDTIYLANVTLPDSPLKVKFKEAAFLPEWVGINADLSPLWTL 362
QY 355 LTP 357
Db 363 AKP 365

```

Search completed: May 27, 2003, 15:06:14  
Job time : 52.1951 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:03:22 ; Search time 15.6237 Seconds  
(without alignments)  
715.625 Million cell updates/sec

Title: us-10-040-863-8

Perfect score: 2032

Sequence: 1 MASAQVFPFPLAHFLIFV.....RSHFLKAKGVTCYVAGRAF 380

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2.6/protdata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2.6/protdata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2.6/protdata/1/iaa/5A\_COMB.pep.\*
- 4: /cgn2.6/protdata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2.6/protdata/1/iaa/6B\_COMB.pep.\*
- 6: /cgn2.6/protdata/1/iaa/6B\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2032	100.0	380	4	US-09-298-886-8
2	1896	93.3	353	4	US-09-298-886-10
3	1502	73.9	340	4	US-09-254-077A-6
4	1502	73.9	340	4	US-09-254-077A-7
5	1485.5	73.1	343	1	US-08-395-800A-8
6	1485.5	73.1	343	4	US-09-254-077A-8
7	1475	72.6	344	4	US-09-298-886-11
8	1472.5	72.5	347	4	US-09-254-077A-9
9	1086.5	53.5	373	4	US-09-254-077A-12
10	1057.5	52.0	365	1	US-07-914-281-6
11	1057.5	52.0	365	1	US-08-393-246-6
12	1057.5	52.0	365	1	US-08-273-411-1
13	1057.5	52.0	365	1	US-08-525-058A-6
14	1057.5	52.0	365	1	US-08-395-800A-6
15	1057.5	52.0	365	1	US-08-395-800A-10
16	1057.5	52.0	365	2	US-08-696-731-6
17	1057.5	52.0	365	4	US-09-042-531-6
18	1057.5	52.0	365	4	US-09-254-077A-11
19	1057.5	52.0	365	5	PC1-US91-00899-12
20	1051.5	51.7	333	5	PC1-US91-00899-11
21	1046.5	51.5	365	4	US-09-151-592-2
22	1046.5	51.5	365	4	US-09-254-077A-10
23	722.5	35.6	222	1	US-08-395-800A-2
24	439	21.6	102	1	US-08-395-800A-3
25	141	6.9	300	4	US-08-433-598-2
26	97	4.8	183	1	US-08-167-035-33
27	97	4.8	183	1	US-08-208-887A-33

28	97	4.8	183	2	US-08-539-005-33	Sequence 33, Appl
29	97	4.8	183	4	US-09-280-598-35	Sequence 35, Appl
30	96.5	4.7	355	4	US-09-390-131-5	Sequence 5, Appl
31	91.5	4.5	801	1	US-07-906-349A-6	Sequence 6, Appl
32	91	4.5	442	1	US-08-220-151-22	Sequence 22, Appl
33	91	4.5	442	1	US-08-413-118-22	Sequence 22, Appl
34	91	4.5	442	3	US-08-473-446-22	Sequence 22, Appl
35	91	4.5	442	6	5470718-3	Patent No. 5470718
36	87	4.3	217	1	US-08-167-035-6	Sequence 6, Appl
37	87	4.3	217	1	US-08-208-887A-6	Sequence 6, Appl
38	87	4.3	217	2	US-08-539-005-6	Sequence 6, Appl
39	87	4.3	217	2	US-08-815-176-3	Sequence 3, Appl
40	87	4.3	217	2	US-08-815-176-3	Sequence 4, Appl
41	87	4.3	217	4	US-08-664-962B-6	Sequence 6, Appl
42	87	4.3	217	4	US-09-311-743-6	Sequence 6, Appl
43	87	4.3	217	4	US-09-280-598-6	Sequence 6, Appl
44	87	4.3	217	4	US-09-197-344-3	Sequence 4, Appl
45	87	4.3	217	4	US-09-197-344-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-09-298-886-8

; Sequence 8, Application US/09298886

; Patent No. 6329170

; GENERAL INFORMATION:

; APPLICANT: Eric H. Holmes et al.

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE

; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES

; FILE REFERENCE: 8511-029

; CURRENT APPLICATION NUMBER: US/09/298,886

; CURRENT FILING DATE: 1999-04-26

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 8

; LENGTH: 380

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-298-886-8

Query Match 100.0%; Score 2032; DB 4; Length 380;  
Best Local Similarity 100.0%; Pred. No. 3.1e-226;  
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MASAQVFPFPLAHFLIFVFTSTIIHQQRIYKQLPLSEKELPMTTOMSSGNTESPEMR	60
Db	1	MASAQVFPFPLAHFLIFVFTSTIIHQQRIYKQLPLSEKELPMTTOMSSGNTESPEMR	60
QY	61	RDSFHGNGELRCMFTINSIGRLGNOMGEYATLALARMNGRLAFIPASMHNALAPIRI	120
Db	61	RDSFHGNGELRCMFTINSIGRLGNOMGEYATLALARMNGRLAFIPASMHNALAPIRI	120
QY	121	SLPVLHSDTAKKIPQWYHNLNDWMEERYRHPGHVFYFTGYSWTYFHLRPEILKEFT	180
Db	121	SLPVLHSDTAKKIPQWYHNLNDWMEERYRHPGHVFYFTGYSWTYFHLRPEILKEFT	180
QY	181	LHDHVRBEAQAFLRGLRVNGSQSTFVGVHVRGDYVHVMPNVKGVVADRGYLEKALDM	240
Db	181	LHDHVRBEAQAFLRGLRVNGSQSTFVGVHVRGDYVHVMPNVKGVVADRGYLEKALDM	240
QY	241	FRARYSPFVFTSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTOCHNTIMTIG	300
Db	241	FRARYSPFVFTSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTOCHNTIMTIG	300
QY	301	TFGIWAAYLAGGTIVILANYTLFDSPLKVKFEAAFLPEWVGIPADLSPLLKALTPACP	360
Db	301	TFGIWAAYLAGGTIVILANYTLFDSPLKVKFEAAFLPEWVGIPADLSPLLKALTPACP	360
QY	361	RSFHHLKAKGVTCYVAGRAF 380	

D6 361 RSHFLKAKGVTCYVAGRAF 380

## RESULT 2

US-09-298-886-10

; Sequence 10, Application US/09298886

; Patent No. 6329170

; GENERAL INFORMATION:

; APPLICANT: Eric H. Holmes et al.

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE

; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 8511-029

; CURRENT FILING DATE: 1999-04-26

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 10

; LENGTH: 353

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-298-886-10

Query Match

Best Local Similarity 93.3%; Score 1896; DB 4; Length 353;

Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LQORIVKLOPSEKELPMTOMSSGNTSPENRDSQHGNGELRGMTINSIGRLGNOM 87

DB 1 LQORIVKLOPSEKELPMTOMSSGNTSPENRDSQHGNGELRGMTINSIGRLGNOM 60

QY 88 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLSHSTAKKIPQNYHLNDWMEER 147

DB 61 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLSHSTAKKIPQNYHLNDWMEER 120

QY 148 YRHIPGHVFTGYPGCSWTFYHHLRPEILKEPTLHDHVREARQAELRGLRVNGSQPSTFV 207

DB 121 YRHIPGHVFTGYPGCSWTFYHHLRPEILKEPTLHDHVREARQAELRGLRVNGSQPSTFV 180

QY 208 GVHVRGQYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVVTSNGMAWRENINAS 267

DB 181 GVHVRGQYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVVTSNGMAWRENINAS 240

QY 268 RGVVVFAGNGIEGSPAKDFALLTQCNTHTMTGTGFIWAAYLAGSDTIYLANITLPDSPF 327

DB 241 RGVVVFAGNGIEGSPAKDFALLTQCNTHTMTGTGFIWAAYLAGSDTIYLANITLPDSPF 300

QY 328 LKVPKPEAAFLPEWVGIPADLSPLKALTPACPRSHFLKAKGVTCYVAGRAF 380

DB 301 LKVPKPEAAFLPEWVGIPADLSPLKALTPACPRSHFLKAKGVTCYVAGRAF 353

## RESULT 3

US-09-254-077A-6

; Sequence 6, Application US/09254077A

; Patent No. 6399758

; GENERAL INFORMATION:

; APPLICANT: SANDRIN, MAURO S.

; APPLICANT: MCKENZIE, IAN C. F.

; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES

; FILE REFERENCE: 30562-SUSWO

; CURRENT FILING DATE: 1999-06-11

; PRIOR FILING DATE: 1997-08-22

; PRIOR FILING DATE: 1997-08-22

; PRIOR FILING DATE: 1996-08-23

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 6

; LENGTH: 340

; TYPE: PRT

; ORGANISM: Sus scrofa

US-09-254-077A-6

Query Match

Best Local Similarity 73.9%; Score 1502; DB 4; Length 340;

Matches 281; Conservative 28; Mismatches 30; Indels 14; Gaps 3;

QY 1 MASAQVPSFFPLAHFLIFVFTSTIIHLQORIVKLOPSEKELPMTOMSSGNTSPENR 60

DB 1 MLMSQASFFFGPFPILFVFTASTIFHLQORIVKLOPSEKELPMTOMSSGNTSPENR 52

QY 61 RDSQHGNGELRGMTINSIGRLGNOMGEYATLALARMNGRLAFIPASMHNALAPIFR 120

DB 53 -----SSPQLKGMWTINAIGRLGNOMGEYATLALARMNGRLAFIPASMHNALAPIFR 106

QY 121 SLPLVLSHSTAKKIPQNYHLNDWMEERYRHIPGHVFTGYPGCSWTFYHHLRPEILKEPT 180

DB 107 TLPVLHASTARRIPQNYHLNDWMEERYRHIPGHVFTGYPGCSWTFYHHLRPEILKEPT 166

QY 181 LHDHVREARQAELRGLRVNGSQPSTFVGVHVRGQYVHVMPNWKGVVADRGYLEKALDM 240

DB 167 LHNHVREARQAELRGLRVNGSQPSTFVGVHVRGQYVHVMPNWKGVVADRGYLEKALDM 226

QY 241 PRARYSSPVVTSNGMAWRENINASRGDVVFAGNGIEGSPAKDFALLTQCNTHTMTIG 300

DB 227 PRARYSSPVVTSNGMAWRENINASRGDVVFAGNGIEGSPAKDFALLTQCNTHTMTIG 286

QY 301 TFGIWAAYLAGSDTIYLANITLPDPSFLKVPKPEAAFLPEWVGIPADLSPLLK 353

DB 287 TFGIWAAYLAGSDTIYLANITLPDPSFLKVPKPEAAFLPEWVGIPADLSPLLK 339

## RESULT 4

US-09-254-077A-7

; Sequence 7, Application US/09254077A

; Patent No. 6399758

; GENERAL INFORMATION:

; APPLICANT: SANDRIN, MAURO S.

; APPLICANT: MCKENZIE, IAN C. F.

; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES

; FILE REFERENCE: 30562-SUSWO

; CURRENT FILING DATE: 1999-06-11

; PRIOR FILING DATE: 1997-08-22

; PRIOR FILING DATE: 1997-08-22

; PRIOR FILING DATE: 1996-08-23

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 7

; LENGTH: 340

; TYPE: PRT

; ORGANISM: Sus scrofa

US-09-254-077A-7

Query Match

Best Local Similarity 73.9%; Score 1502; DB 4; Length 340;

Matches 281; Conservative 28; Mismatches 30; Indels 14; Gaps 3;

QY 1 MASAQVPSFFPLAHFLIFVFTSTIIHLQORIVKLOPSEKELPMTOMSSGNTSPENR 60

DB 1 MLMSQASFFFGPFPILFVFTASTIFHLQORIVKLOPSEKELPMTOMSSGNTSPENR 52

QY 61 RDSQHGNGELRGMTINSIGRLGNOMGEYATLALARMNGRLAFIPASMHNALAPIFR 120

DB 53 -----SSPQLKGMWTINAIGRLGNOMGEYATLALARMNGRLAFIPASMHNALAPIFR 106

QY 121 SLPLVLSHSTAKKIPQNYHLNDWMEERYRHIPGHVFTGYPGCSWTFYHHLRPEILKEPT 180

DB 107 TLPVLHASTARRIPQNYHLNDWMEERYRHIPGHVFTGYPGCSWTFYHHLRPEILKEPT 166

QY 181 LHDHVREARQAELRGLRVNGSQPSTFVGVHVRGQYVHVMPNWKGVVADRGYLEKALDM 240

DB 167 LHNHVREARQAELRGLRVNGSQPSTFVGVHVRGQYVHVMPNWKGVVADRGYLEKALDM 226





```

US-07-914-281-6
; Sequence 6, Application US/07914281
; Patent No. 5324663
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOPOLYMER, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07914,281
; FILING DATE: 19920720
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 6:
; LENGTH: 365 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-07-914-281-6

Query Match 52.0%; Score 1057.5; DB 1; Length 365;
Best Local Similarity 56.5%; Pred. No. 1.5e-113;
Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps 4;

QY 12 LAHFLIFVFVSTIIHQ-----RIVKLQPLSEKELPMTTOMSSGNTE 56
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 11 LAFLVLCVLSVIFFLHIDQSPFPHGLSLCPLDRLVTPPVAIFCLPGTAMGNASSSC 70
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 57 PEMRRDSQHGNGELRGFTINSTIGRLGNOMGEYATILFALARMNGRLAIFPASMHNALAP 116
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 P-----QH-PASLSGTWTVPNGRGNOMGQYATLLAQLNRRRAFILPAMEAALAP 122
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 117 IFRSLPVLHSDTAKKIPWQNYHLNDWMEERYRHPGFVFTGPGSWTFYHHLRPEIL 176
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 123 VFRITPLVLAPEVDSRTPWRLEQLHDMWSEYADLRDPFLKLSGFPSCSWTFHHLRQIR 182
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 177 KEFTLHDHVRREAQAFILRGLV--NGSQPSTFVGVHVRGQYVHVMNVKGVVADRGYL 234
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 183 REFTLHDHVRREAQAFILRGLV--NGSQPSTFVGVHVRGQYVHVMNVKGVVADRGYL 242
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 235 EKALDMFRARYSSPVFVVTSSNGMAWCRENINASRGDVVVFAGNGIEGSPAKDFALLITQCNH 294
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 243 RQADWFRARAEAPVFTVSSNGMEWCKENIDTSQDVTTFAGDQGEAPLWKDFALLITQCNH 302
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 295 TIMTIGTGIWAAYLAGDGTIYLANITLPSDFLTKVRKPEAAFLPEWYGPADLSPILKA 354
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 303 TIMTIGTGIWAAYLAGDGTIYLANITLPSDFLTKVRKPEAAFLPEWYGPADLSPILKA 362
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 355 LTP 357
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```
Db 243 ROAMDWFRARHEAPVFTVTSNGMCKENIDTSQGDVTFAGDQCEATPWKDFALLTQC�H 302
QY 295 TIMTIGTGGWAAAYLAGGDTIYLANVTLDPSPFLKVKPEAAFLPWVGIPADLSPLLKA 354
Db 303 TIMTIGTGGWAAAYLAGGDTIYLANVTLDPSPFLKVKPEAAFLPWVGIPADLSPLLKA 362
QY 355 LTP 357
Db 363 AKP 365

RESULT 12
US-08-273-411-1
: Sequence 1, Application US/08273411
: Patent No. 5625124
: GENERAL INFORMATION:
: APPLICANT: Falk, Per
: APPLICANT: Gordon, Jeffrey I.
: TITLE OF INVENTION: Animal Model for Gastro-Intestinal
: TITLE OF INVENTION: Disease
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Patrea L. Pabst
: STREET: 1100 Peachtree Street, Suite 2800
: CITY: Atlanta
: STATE: Georgia
: COUNTRY: USA
: ZIP: 30309-4530
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/273,411
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Pabst, Patrea L.
: REGISTRATION NUMBER: 31,284
: REFERENCE/DOCKET NUMBER: WU106
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (404) 815-6508
: TELEFAX: (404) 815-6555
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 365 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE: internal
: FEATURE:
: NAME/KEY: misc-feature
: LOCATION: 1..365
: OTHER INFORMATION: /note= "GDP-L-fucose:beta-D-Galactoside-2-alpha-L-fucosyl
: PUBLICATION INFORMATION:
: AUTHORS: Larsen, et al.
: JOURNAL: Proc. Nat'l Acad. Sci. USA
: VOLUME: 87
: PAGES: 6674-6678
: DATE: 1990
: RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 365
US-08-273-411-1

Query Match 52.0%; Score 1057.5; DB 1; Length 365;
Best Local Similarity 56.5%; Pred. No. 1.5e-113;
Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps 4;
QY 12 LAEFLFVFTVTTIHLQ-----RIVKLQPLSEKELPMTTQMSSGNIES 56
```

```
Db 11 LAEFLVCVLSVIFFLIHODSPFHGLGLSTLQDRLVTPPPVAIFCLPGTAMGNASSSC 70
QY 57 PEMRRSEOHNGELRGMTTINSIGRLGNOMGYAIFALARMNGRLAIFAPSMHNALAP 116
Db 71 P-----QH-PASLSGTIVTPNGRFGNQMGQYATLLAALQINGRRATFLFAMHAALAP 122
QY 117 IFRISLPVLHSDTAKKIPWQNYHLNDWMEERYHIFGHEVFRFTGYPCSWTEYHHLRPEIL 176
Db 123 VFRITLPVLAPEVDSTFPRRELQLDHDMSEYADLRDPFLKLSGFCPSWTFPHHLREQIR 182
QY 177 KEFTLHDHVREEAQAFRLGLRV--NGSQPSTFVGVHVRKGDYVHVMPNWKGVVADRGYL 234
Db 183 REFTLHDHLREEAQSGLQGLRLGRTGDRPTFVGHVHRKGDYLVQNVMPQWKGVGVGSAYL 242
QY 235 EKALDMFRARYSSPVFVTSNGMACKRENINASRGDVTAGNIGESPAKDFALLTQC�H 294
Db 243 ROAMDWFRARHEAPVFTVTSNGMCKENIDTSQGDVTFAGDQCEATPWKDFALLTQC�H 302
QY 295 TIMTIGTGGWAAAYLAGGDTIYLANVTLDPSPFLKVKPEAAFLPWVGIPADLSPLLKA 354
Db 303 TIMTIGTGGWAAAYLAGGDTIYLANVTLDPSPFLKVKPEAAFLPWVGIPADLSPLLKA 362
QY 355 LTP 357
Db 363 AKP 365

RESULT 13
US-08-525-058A-6
: Sequence 6, Application US/08525058A
: Patent No. 5770420
: GENERAL INFORMATION:
: APPLICANT: LOWE, JOHN B.
: TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
: TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
: TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT, P.C.
: STREET: 1755 Jefferson Davis Highway, Fourth Floor
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/525,058A
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Lavalleye, Jean-Paul M. P.
: REGISTRATION NUMBER: 31,451
: REFERENCE/DOCKET NUMBER: 2363-060-55
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 521-4500
: TELEFAX: (703) 486-2347
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 365 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-525-058A-6

Query Match 52.0%; Score 1057.5; DB 1; Length 365;
Best Local Similarity 56.5%; Pred. No. 1.5e-113;
```







GenCore version: 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:08:29 ; Search time 16.4181 Seconds  
(without alignments)  
2295.367 Million cell updates/sec

Title: US-10-040-863-8

Perfect score: 2032

Sequence: 1 MASAQVPFSPPLAHFLIFVE.....RSHFLHAKAGTCVYAGRAF 380

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA.\*

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pap.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PTI\_NEW\_PUB.pap.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pap.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pap.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pap.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pap.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PTUS\_PUBCOMB.pap.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pap.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pap.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pap.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pap.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pap.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pap.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2032	100.0	380	10	US-09-999-672-8
2	2032	100.0	380	12	US-10-040-863-8
3	1896	93.3	353	10	US-09-999-672-10
4	1896	93.3	353	12	US-10-040-863-10
5	1502	72.9	340	10	US-09-051-034A-2
6	1475	72.6	344	10	US-09-999-672-11
7	1475	72.6	344	12	US-10-040-863-11
8	1057.5	52.0	365	9	US-10-105-963-10
9	1057.5	52.0	365	10	US-09-863-475A-6
10	1046.5	51.5	365	10	US-09-051-034A-4
11	1044.5	51.4	365	10	US-09-844-268-13
12	1044.5	51.4	365	10	US-09-844-705-13
13	141	6.9	300	10	US-09-848-838-2
14	113	5.6	431	9	US-09-479-614-14
15	113	5.6	496	9	US-09-479-614-2
16	110.5	5.4	496	9	US-10-214-524-25
17	109	5.4	496	9	US-09-479-614-29
18	89	4.4	1031	9	US-09-764-868-647
19	89	4.4	2057	9	US-09-815-379-10

20	89	4.4	2058	9	US-09-815-379-17	Sequence 17, Appl
21	87	4.3	217	10	US-09-765-298A-6	Sequence 6, Appl
22	85.5	4.2	738	9	US-10-235-521-1	Sequence 1, Appl
23	84.5	4.2	344	9	US-09-895-913A-56	Sequence 56, Appl
24	84.5	4.2	344	10	US-09-815-242-11366	Sequence 11366, A
25	83.5	4.1	747	9	US-10-028-072-212	Sequence 212, App
26	83.5	4.1	747	9	US-10-121-049-212	Sequence 212, App
27	83.5	4.1	747	9	US-10-123-904-212	Sequence 212, App
28	83.5	4.1	747	9	US-10-140-470-212	Sequence 212, App
29	83.5	4.1	747	9	US-10-175-746-212	Sequence 212, App
30	83.5	4.1	747	9	US-10-176-918-212	Sequence 212, App
31	83.5	4.1	747	9	US-10-176-921-212	Sequence 212, App
32	83.5	4.1	747	9	US-10-137-865-212	Sequence 212, App
33	83.5	4.1	747	9	US-10-140-474-212	Sequence 212, App
34	83.5	4.1	747	9	US-10-142-431-212	Sequence 212, App
35	83.5	4.1	747	9	US-10-143-114-212	Sequence 212, App
36	83.5	4.1	747	9	US-10-140-002-212	Sequence 212, App
37	83.5	4.1	747	9	US-10-142-419-212	Sequence 212, App
38	83.5	4.1	747	9	US-10-123-262-212	Sequence 212, App
39	83.5	4.1	747	9	US-10-142-423-212	Sequence 212, App
40	83.5	4.1	747	9	US-10-121-050-212	Sequence 212, App
41	83.5	4.1	747	9	US-10-141-755-212	Sequence 212, App
42	83.5	4.1	747	9	US-10-143-032-212	Sequence 212, App
43	83.5	4.1	747	9	US-10-123-108-212	Sequence 212, App
44	83.5	4.1	747	9	US-10-123-236-212	Sequence 212, App
45	83.5	4.1	747	9	US-10-123-261-212	Sequence 212, App

#### ALIGNMENTS

RESULT 1

US-09-999-672-8

; Sequence 8, Application US/09999672

; Patent No. US20020127655A1

; GENERAL INFORMATION:

; APPLICANT: Eric H. Holmes et al.

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE

; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES

; FILE REFERENCE: 8511-029

; CURRENT APPLICATION NUMBER: US/09/999,672

; PRIOR FILING DATE: 2001-10-31

; PRIOR APPLICATION NUMBER: US/09/298,886

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 8

; LENGTH: 380

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-999-672-8

Query Match 100.0%; Score 2032; DB 10; Length 380;  
Best Local Similarity 100.0%; Pred. No. 1.6e-200;  
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MASAQVPFSPPLAHFLIFVFTSTIIHQORIVKLOPLSEKELPMTQSSNGTSPEMR	60
DB	1	MASAQVPFSPPLAHFLIFVFTSTIIHQORIVKLOPLSEKELPMTQSSNGTSPEMR	60
QY	61	RSEQHNGELRGMTTINSIGRLGNOMGEYATLALARNNGRLAFIPASMHNALAFIPRI	120
DB	61	RSEQHNGELRGMTTINSIGRLGNOMGEYATLALARNNGRLAFIPASMHNALAFIPRI	120
QY	121	SLPVLHSDTAKKIPQNYHLNDWMEERYRHIFGHVFRFTGYPCSWTFYHHLPEILKERT	180
DB	121	SLPVLHSDTAKKIPQNYHLNDWMEERYRHIFGHVFRFTGYPCSWTFYHHLPEILKERT	180
QY	181	LHDHYREEAAQALRGLRVNGSQPSTFVGHVRRGDYVHVMPNVNKGVADRGYLEKALDM	240
DB	181	LHDHYREEAAQALRGLRVNGSQPSTFVGHVRRGDYVHVMPNVNKGVADRGYLEKALDM	240

```

QY 241 FRARYSSPVVVTSSNGMAWCENINASRGDVVFAGNGIEGSPAKDFALLTQCNTHTMTIG 300
|||||
Db 241 FRARYSSPVVVTSSNGMAWCENINASRGDVVFAGNGIEGSPAKDFALLTQCNTHTMTIG 300
|||||
QY 301 TFGIWAAYLAGDIIYLANYTLPDSEFLKVEKPEAAFLPEWVGIPADLSPLLKALTPACP 360
|||||
Db 301 TFGIWAAYLAGDIIYLANYTLPDSEFLKVEKPEAAFLPEWVGIPADLSPLLKALTPACP 360
|||||
QY 361 RSHFHLKAKGVTCYVAGRAF 380
|||||
Db 361 RSHFHLKAKGVTCYVAGRAF 380
|||||

RESULT 2
US-10-040-863-8
; Sequence 8, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/298,886
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-040-863-8

Query Match 100.0%; Score 2032; DB 12; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.6e-200;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAQVPSFPLAHFLIEFVVTSTIIHLQQRIVKLOPLSEKELPMTTQSSNGTSPEMR 60
|||||
Db 1 MASAQVPSFPLAHFLIEFVVTSTIIHLQQRIVKLOPLSEKELPMTTQSSNGTSPEMR 60
|||||
QY 61 RDSQHGNGELRGMTINSIGRLNQMGYEATLALARMNGRLAFIPASMHNALAPIFRI 120
|||||
Db 61 RDSQHGNGELRGMTINSIGRLNQMGYEATLALARMNGRLAFIPASMHNALAPIFRI 120
|||||
QY 121 SLPLVLSHTAKKIPWNYHLNDWMEERYHHPGHEVFTGYPCSWTFYHHLRPEILKKEFT 180
|||||
Db 121 SLPLVLSHTAKKIPWNYHLNDWMEERYHHPGHEVFTGYPCSWTFYHHLRPEILKKEFT 180
|||||
QY 181 LHDHVREAAQFLRGLRVNGSQPSTFVGHVHVRGDDYVHVPNVKGVVADRGYLEKALDM 240
|||||
Db 181 LHDHVREAAQFLRGLRVNGSQPSTFVGHVHVRGDDYVHVPNVKGVVADRGYLEKALDM 240
|||||
QY 241 FRARYSSPVVVTSSNGMAWCENINASRGDVVFAGNGIEGSPAKDFALLTQCNTHTMTIG 300
|||||
Db 241 FRARYSSPVVVTSSNGMAWCENINASRGDVVFAGNGIEGSPAKDFALLTQCNTHTMTIG 300
|||||
QY 301 TFGIWAAYLAGDIIYLANYTLPDSEFLKVEKPEAAFLPEWVGIPADLSPLLKALTPACP 360
|||||
Db 301 TFGIWAAYLAGDIIYLANYTLPDSEFLKVEKPEAAFLPEWVGIPADLSPLLKALTPACP 360
|||||
QY 361 RSHFHLKAKGVTCYVAGRAF 380
|||||
Db 361 RSHFHLKAKGVTCYVAGRAF 380
|||||

RESULT 3
US-09-999-672-10
; Sequence 10, Application US/09999672
; Patent No. US20020127655A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/999,672
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US/09/298,886
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-999-672-10

Query Match 93.3%; Score 1896; DB 10; Length 353;
Best Local Similarity 100.0%; Pred. No. 1.4e-186;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LQQRIVKLOPLSEKELPMTTQSSNGTSPENRRDSEQHGNGELRGMTINSIGRLGNOM 87
|||||
Db 1 LQQRIVKLOPLSEKELPMTTQSSNGTSPENRRDSEQHGNGELRGMTINSIGRLGNOM 60
|||||
QY 88 GEYATLALARMNGRLAFIPASMHNALAPIFRIISLPVLSHDTAKKIPWNYHLNDWMEER 147
|||||
Db 61 GEYATLALARMNGRLAFIPASMHNALAPIFRIISLPVLSHDTAKKIPWNYHLNDWMEER 120
|||||
QY 148 YRHIPGHEVFRFTGYPCSWTFYHHLRPEILKKEFTLHDHVREAAQFLRGLRVNGSQPSTFV 207
|||||
Db 121 YRHIPGHEVFRFTGYPCSWTFYHHLRPEILKKEFTLHDHVREAAQFLRGLRVNGSQPSTFV 180
|||||
QY 208 GYHVRGDDYVHVPNVKGVVADRGYLEKALDMFRARYSSPVVVTSSNGMAWCENINAS 267
|||||
Db 181 GYHVRGDDYVHVPNVKGVVADRGYLEKALDMFRARYSSPVVVTSSNGMAWCENINAS 240
|||||
QY 268 RGVVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTGIWAAYLAGDIIYLANYTLPDSPF 327
|||||
Db 241 RGVVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTGIWAAYLAGDIIYLANYTLPDSPF 300
|||||
QY 328 LKVFKEAAFLPEWVGIPADLSPLLKALTPACPRSHFHLKAKGVTCYVAGRAF 380
|||||
Db 301 LKVFKEAAFLPEWVGIPADLSPLLKALTPACPRSHFHLKAKGVTCYVAGRAF 353
|||||

RESULT 4
US-10-040-863-10
; Sequence 10, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/298,886
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-040-863-10

Query Match 93.3%; Score 1896; DB 12; Length 353;
Best Local Similarity 100.0%; Pred. No. 1.4e-186;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LQQRIVKLOPLSEKELPMTTQSSNGTSPENRRDSEQHGNGELRGMTINSIGRLGNOM 87

```

Db 1 LQORIVKLOPLSEKELPMTQSSNGELSPERDSEHGNGELRGFMFINSIGRLGNQM 60  
QY 88 GEYATLAFALARMNGRLAFIPASMHNALAFIPRISLVLSDTAKKIPWQNYHLNDWMEER 147  
Db 61 GEYATLAFALARMNGRLAFIPASMHNALAFIPRISLVLSDTAKKIPWQNYHLNDWMEER 120  
QY 148 YRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGRLVNGSQSPSTFV 207  
Db 121 YRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGRLVNGSQSPSTFV 180  
QY 208 GVHVRGDIYVHVMNPVWKGVDARGYLEKALDMFRARYSSPVFVTSNGMAWCRENINAS 267  
Db 181 GVHVRGDIYVHVMNPVWKGVDARGYLEKALDMFRARYSSPVFVTSNGMAWCRENINAS 240  
QY 268 RGVDFVAGNGIEGSPAKDFALLTQCNTMTIGTFGIWAAYLAGGDTIYLANITLDPSPF 327  
Db 241 RGVDFVAGNGIEGSPAKDFALLTQCNTMTIGTFGIWAAYLAGGDTIYLANITLDPSPF 300  
QY 328 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 380  
Db 301 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 353

## RESULT 5

US-09-051-034A-2  
; Sequence 2, Application US/09051034A  
; Patent No. US2001005584A1  
; GENERAL INFORMATION:

; APPLICANT: MCKENZIE, IAN FARQUHAR CAMPBELL  
; APPLICANT: SANDRIN, MAURO SERGIO  
; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC  
; TITLE OF INVENTION: GLYCOSYLTRANSFERASE  
; FILE REFERENCE: 30562.6USWO  
; CURRENT APPLICATION NUMBER: US/09/051.034A  
; CURRENT FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: PCT/AU97/00492  
; PRIOR FILING DATE: 1997-08-01  
; PRIOR APPLICATION NUMBER: 60/024,279  
; PRIOR FILING DATE: 1996-08-21  
; PRIOR APPLICATION NUMBER: P01402  
; PRIOR FILING DATE: 1996-08-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 340  
; TYPE: PRT  
; ORGANISM: Sus Domesticus

## US-09-051-034A-2

Query Match 73.9%; Score 1502; DB 10; Length 340;  
Best Local Similarity 79.6%; Pred. No. 4.5e-146;  
Matches 281; Conservative 28; Mismatches 30; Indels 14; Gaps 3;

QY 1 MASAOVPFSPFLAHFLIFVFTVSTIIHLQORIVKLOPLSEKELPMTQSSNGTESPEMR 60  
Db 1 MLSMOASFPFPGPFLFVETASTIFHLQORVVKIQP--TWELQVTVT---TESP--- 52  
QY 61 RSEHQHNGELGDMFTINSIGRLGNOMGEYATLAFALARMNGRLAFIPASMHNALAFIPRI 120  
Db 53 -----SSPOLKGMWTINAIIGRLGNOMGEYATLAFALARMNGRPAFIPPMHSTLAFIPRI 106  
QY 121 SLPLVHSDTAKKIPWQNYHLNDWMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFT 180  
Db 107 TLPLVHSTARIIPWQNYHLNDWMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFT 166  
QY 181 LHDHVREAAQAFRLGRLVNGSQSPSTFVGVHVRGDIYVHVMNPVWKGVDARGYLEKALDM 240  
Db 167 LHNHVREAAQDFLRLGRLVNGSRPSTVGVHVRGDIYVHVMNPVWKGVDARYLEQALDM 226  
QY 241 FRARYSSPVFVTSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTQCNTMTIG 300  
Db 227 FRARYSSPVFVTSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTQCNTMTIG 286

QY 301 TFGIWAAYLAGGDTIYLANITLDPSPFLKVKPEAAFLPEWVGIPADLSPLLK 353  
Db 287 TFGIWAAYLAGGDTIYLANITLDPSPFLKVKPEAAFLPEWVGIPADLSPLLK 339  
RESULT 6  
US-09-999-672-11  
; Sequence 11, Application US/09999672  
; Patent No. US20020127655A1  
; GENERAL INFORMATION:  
; APPLICANT: ERIC H. HOLMES ET AL.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; TITLE OF INVENTION: GMI-SPECIFIC ALPHAI-2 FUCOSYLTRANSFERASE AND USES  
; FILE REFERENCE: 8511-029  
; CURRENT APPLICATION NUMBER: US/09/999,672  
; CURRENT FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: US/09/298,886  
; PRIOR FILING DATE: 1999-04-26  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 344  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-999-672-11

Query Match 72.6%; Score 1475; DB 10; Length 344;  
Best Local Similarity 77.7%; Pred. No. 2.7e-143;  
Matches 275; Conservative 30; Mismatches 37; Indels 12; Gaps 3;

QY 1 MASAOVPFSPFLAHFLIFVFTVSTIIHLQORIVKLOPLSEKELPMTQSSNGTESPEMR 60  
Db 1 MLVYQMPFSPFMAHFLIFVFTVSTIFHVQORLAKIQAM--WELPV-----QIPVLA 49  
QY 61 RSEHQHNGELGDMFTINSIGRLGNOMGEYATLAFALARMNGRLAFIPASMHNALAFIPRI 120  
Db 50 STSKALGPSOLRCGMWTINAIIGRLGNOMGEYATLAFALARMNGRPAFIPQMHSTLAFIPRI 109  
QY 121 SLPLVHSDTAKKIPWQNYHLNDWMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFT 179  
Db 110 TLPLVHSTARIIPWQNYHLNDWMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFT 169  
QY 180 TLHDHVREAAQAFRLGRLVNGSQSPSTFVGVHVRGDIYVHVMNPVWKGVDARGYLEKALD 239  
Db 170 TLHDHVREAAQAFRLGRLVNGSRPSTFVGVHVRGDIYVHVMNPVWKGVDARYLQOALD 229  
QY 240 MFRARYSSPVFVTSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTQCNTMTIG 299  
Db 230 WFEARYSSLIFFVTSNGMAWCRENIDTSHGDVVFAGNGIEGSPAKDFALLTQCNTMTIG 289  
QY 300 TFGIWAAYLAGGDTIYLANITLDPSPFLKVKPEAAFLPEWVGIPADLSPLLK 353  
Db 290 TFGIWAAYLAGGDTIYLANITLDPSPFLKVKPEAAFLPEWVGIPADLSPLLK 343

## RESULT 7

US-10-040-863-11  
; Sequence 11, Application US/10040863  
; Patent No. US20020137165A1  
; GENERAL INFORMATION:

; APPLICANT: ERIC H. HOLMES ET AL.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; TITLE OF INVENTION: GMI-SPECIFIC ALPHAI-2 FUCOSYLTRANSFERASE AND USES  
; FILE REFERENCE: 8511-029  
; CURRENT APPLICATION NUMBER: US/10/040,863  
; CURRENT FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: 09/298,886  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0

OV 12 LABELFVFVSTIHL00-----RTVKLOPI.SEKEL.PMTTOMSSGNTE\$ 56

Db 11 LAFLVGVLSVIFELHHDSPFPHGLGLSILCPDRRLVTPPVAFICLPGTAMGNSASSC 70  
Qy 57 PEMRSDSEHGNGELRGMTINSIGRLGNOMGEYATLALARMNGRLAFIPASMHNLAP 116  
Db 71 P-----QH-PASLSTWTVYPNGRGNQMGQYATLALAOQLNRAFIPLPAMHAALAP 122  
Qy 117 IFRISLPLVHSDTAKKIPWQNYHLNDWMEERYRHIPGHEVFRFTGYPCSWTFYHHLRPEIL 176  
Db 123 VFRITLVLAPENDSRTPWRLOLHDWMSSEYADLRDPFLKLSOFFPCSWTFYHHLRQIR 182  
Qy 177 KEFTLHDHVBREAQAFIRGRV--NGSQPSTFVGVHVRGDDYVHVMNVKGVVADRGYL 234  
Db 183 REFTLHDHVBREAQSVLQRLGRIGDRPRTFVGVHVRGDDYLVQMPQRMKGVVGD SAYL 242  
Qy 235 EKALDMPFARYSSVFFVYTSNGMAWCARENINASRGDVFVAGNGIEGSPAKDFALLTQCNI 294  
Db 243 RQAMDWFRARAEAFVFFVYTSNGMECKENIDTSGDVTFFAGDGEATPWKDFALLTQCNI 302  
Qy 295 TIMTIGFGIWAAYLAGDITIYLANFTLPDSFPLKVFKEPAAFPEWVGIPADLSPLIKA 354  
Db 303 TIMTIGFGIWAAYLAGDITVYLANFTLPDSFPLKVFKEPAAFPEWVGINADLSPLWTL 362  
Qy 355 LTP 357  
Db 363 AKP 365

## RESULT 10

US-09-051-034A-4  
; Sequence 4, Application US/09051034A  
; Patent No. US20010055584A1  
; GENERAL INFORMATION:  
; APPLICANT: MCKENZIE, IAN FAROUKH CAMPBELL  
; APPLICANT: SANDRIN, MAURO SERGIO  
; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC  
; FILE REFERENCE: GLYCOSYLTRANSFERASE  
; FILE REFERENCE: 30562.6USWO  
; CURRENT APPLICATION NUMBER: US/09/051.034A  
; CURRENT FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: PCT/AU97/00492  
; PRIOR FILING DATE: 1997-08-01  
; PRIOR APPLICATION NUMBER: 60/024,279  
; PRIOR FILING DATE: 1996-08-21  
; PRIOR APPLICATION NUMBER: P01402  
; PRIOR FILING DATE: 1996-08-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 365  
; TYPE: PRT  
; ORGANISM: Sus Domesticus  
US-09-051-034A-4

Query Match 51.5%; Score 1046.5; DB 10; Length 365;  
Best Local Similarity 62.6%; Pred. No. 3.6e-99;  
Matches 194; Conservative 38; Mismatches 75; Indels 3; Gaps 2;

Qy 44 PMTQMSSGNTESPEMRDSEHQHNGELRGMTINSIGRLGNOMGEYATLALARMNGRL 103  
Db 51 PVAIFCLAGTVPVHPNADSDCPKH-PASESGTWITIPDGRFGNQMGQYATLALAOQLNGRQ 109  
Qy 104 AFIPASMHNLAPIFRISLPLVHSDTAKKIPWQNYHLNDWMEERYRHIPGHEVFRFTGYPC 163  
Db 110 AFIQPAMHAVLAPVERITLPLVLAPEVDHRAWRELELDHMSSEYAHLEKPEWLKLTGPPC 169  
Qy 164 SWTFYHHLRPEILKEFTLHDHVBREAQAFIRGRV--NGSQPSTFVGVHVRGDDYVHVMNV 221  
Db 170 SWTFYHHLRPEIRSEFTLHDHVBREAQAFIRGRV--NGSQPSTFVGVHVRGDDYVHVMNV 229  
Qy 222 NWKGVVADRGYLEKALDMFRARYSSPVVYTSNGMAWCARENINASRGDVFVAGNGIEGS 281  
Db 230 KRWKGVVGDGYLOQAMDWFARYEAPVFFVYTSNGMECKENIDTSGDVTFFAGDGEATPWKDFALLTQCNI 289

Qy 282 PAKDFALLTQCNIHTIMTIGTFCIWAAYLAGDITIYLANFTLPDSFPLKVFKEPAAFLEW 341  
Db 290 PARDFALLVQCNIHTIMTIGTFCIWAAYLAGDITIYLANFTLPDSFPLKVFKEPAAFLEW 349  
Qy 342 VGIPADLSPL 351  
Db 350 VGINADLSPL 359

## RESULT 11

US-09-844-268-13  
; Sequence 13, Application US/09844268  
; Patent No. US20020129395A1  
; GENERAL INFORMATION:  
; APPLICANT: BOSWORTH, BRAD  
; APPLICANT: VOGELI, PETER  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY  
; FILE REFERENCE: 21419/90368  
; CURRENT APPLICATION NUMBER: US/09/844.268  
; CURRENT FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: 09/443,766  
; PRIOR FILING DATE: 1999-11-19  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 365  
; TYPE: PRT  
; ORGANISM: Porcine  
US-09-844-268-13

Query Match 51.4%; Score 1044.5; DB 10; Length 365;  
Best Local Similarity 62.6%; Pred. No. 5.8e-99;  
Matches 194; Conservative 38; Mismatches 75; Indels 3; Gaps 2;

Qy 44 PMTQMSSGNTESPEMRDSEHQHNGELRGMTINSIGRLGNOMGEYATLALARMNGRL 103  
Db 51 PVAIFCLAGTVPVHPNADSDCPKH-PASESGTWITIPDGRFGNQMGQYATLALAOQLNGRQ 109  
Qy 104 AFIPASMHNLAPIFRISLPLVHSDTAKKIPWQNYHLNDWMEERYRHIPGHEVFRFTGYPC 163  
Db 110 AFIQPAMHAVLAPVERITLPLVLAPEVDHRAWRELELDHMSSEYAHLEKPEWLKLTGPPC 169  
Qy 164 SWTFYHHLRPEILKEFTLHDHVBREAQAFIRGRV--NGSQPSTFVGVHVRGDDYVHVMNV 221  
Db 170 SWTFYHHLRPEIRSEFTLHDHVBREAQAFIRGRV--NGSQPSTFVGVHVRGDDYVHVMNV 229  
Qy 222 NWKGVVADRGYLEKALDMFRARYSSPVVYTSNGMAWCARENINASRGDVFVAGNGIEGS 281  
Db 230 KRWKGVVGDGYLOQAMDWFARYEAPVFFVYTSNGMECKENIDTSGDVTFFAGDGEATPWKDFALLTQCNI 289  
Qy 282 PAKDFALLTQCNIHTIMTIGTFCIWAAYLAGDITIYLANFTLPDSFPLKVFKEPAAFLEW 341  
Db 290 PARDFALLVQCNIHTIMTIGTFCIWAAYLAGDITIYLANFTLPDSFPLKVFKEPAAFLEW 349  
Qy 342 VGIPADLSPL 351  
Db 350 VGINADLSPL 359

## RESULT 12

US-09-844-705-13  
; Sequence 13, Application US/09844705  
; Patent No. US20020133836A1  
; GENERAL INFORMATION:  
; APPLICANT: BOSWORTH, BRAD  
; APPLICANT: VOGELI, PETER  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY  
; FILE REFERENCE: 21419/90368  
; CURRENT APPLICATION NUMBER: US/09/844.705  
; CURRENT FILING DATE: 2001-04-27

```

; PRIOR APPLICATION NUMBER: 09/443,766
;
; PRIOR FILING DATE: 1999-11-19
;
; NUMBER OF SEQ ID NOS: 13
;
; SOFTWARE: Patentin Ver. 2.0
;
; SEQ ID NO 13
;
; LENGTH: 365
;
; TYPE: PRT
;
; ORGANISM: porcine
;
; US-09-844-705-13

```

Query Match	51.4%;	Score 1044.5;	DB 10;	Length 365;
Best Local Similarity	62.6%;	Pred. No. 5.8e-99;		
Matches 194;	Conservative 38;	Mismatches 75;	Indels 3;	

QY	44	PNTQMSNGNTSP	EMRRDSEQHGNGELRMFTINSIGRLGNQMGYATIFALARMNGRL	103
		:   :	:   :	
Db	51	PAVFAFLCGLTPV	PNASDSCPKH-PASFSGTITYPDGRGNGMGYATLLAQLNGRQ	109
		:   :	:   :	
QY	104	AFIPASMHNALAPI	FRISLPVLHSDTAKIPWNYHLNDWMEERYRHPGHFVRTGYPC	163
		:   :	:   :	
Db	110	AFIQPAMHVA	LAPVPRITLPVLAPEVDYRHAPWELEHLDWMSDYAHLKEPWLKLGFFC	169
		:   :	:   :	
QY	164	SWTFYHLHLPETIL	KEPTLHDHVRERAQAFILRLRV-NGSQSTFVGVHVRGDDYVHVP	221
		:    :	:    :	
Db	170	SWTFPHHLREQIR	SEPTLHDHVRERAQAFILRLRV-NGSQSTFVGVHVRGDDYVHVP	229
		:    :	:    :	
QY	222	NVWKGVVADRGYLE	KALDMFRARYSPFVVTNSGMACRENTNAGSGDVFVAGNGIEG	281
		:    :	:    :	
Db	230	KRWKGVVGDGRYL	QQAMDWFARYEAPFVVTNSGMENCRKKNDITSRGDIVIAGDGREAA	289
		:    :	:    :	
QY	282	PAKDFALLTQCNT	INTTGTGFWAAYLAGGDTIYLANVLPDPSFLKVKFPEEAAFLPEW	341
		:    :	:    :	
Db	290	PAKDFALLVOCNT	INTTGTGFWAAYLAGGDTIYLANVLPDPSFLKVKFPEEAAFLPEW	349
		:    :	:    :	
QY	342	VGIPADLSPL	351	
		:    :	:    :	
Db	350	VGINADLSPL	359	
		:    :	:    :	

```

RESULT 13
US-09-848-838-2
; Sequence 2, Application US/09848838
; Patent No. US2002003750A1
; GENERAL INFORMATION:
; APPLICANT: Taylor, Diane
; APPLICANT: Wang, Ge
; APPLICANT: Falciuc, Monica
; TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
; FILE REFERENCE: 07254-061002
; CURRENT APPLICATION NUMBER: US/09/848,838
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: US 09/433,598
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: US 60/107,268
; PRIOR FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PR1
; ORGANISM: Helicobacter pylori
US-09-848-838-2

```

Query Match 6.9%; Score 141; DB 10; Length 300;  
Best Local Similarity 24.0%; Pred. No. 3.Be-06;  
Matches 81; Conservative 33; Mismatches 125; Indels 98; Gaps 14;

```
QY      81 GRIGNOMGEYA-----TLFALARNGRILAFPSWHNALAPFRIS-- 121
       ||||| :||
Db      10 GGLGNMFOYAFAKSLQKHLNTPVLLDTTSFDNSNRKMQLFPIDLPYANAKEIAKM 69
       ||||| :||
QY     122 --LPLVHSTAKIPQNYHLNDWMEEVTHI--PGHFVFRTGPCSWTYYHHLRPEILKE 178
       ||||| :||
```

```

Db 70 QHLPKLVDAKLYIGFDRVSOEIVFEYEPXLLKPSRLTYFFGYFQDFRYFDAISSLIKOT 129
Qy 179 FTL-----HCHVREEAQAFGLRVNGSQSPSTFVGHVHVRGDDVHVMPVWKGSVTA 229
Db 130 PTLPEPPENKNNKKEEYQRKLS--LILAANSVF--VHIERGYVGI-----GCOL 179
Qy 230 DRGYLEXALDMFRARYSPFVVTYSGMAMCRNTINASRGDVVVFAGNGIEGSPAKDFA-- 287
Db 180 GIDYOKKALEYNAKRVNMBELV-----PCE-----DLKFTQNLQILGYPTDFTKTR 225
Qy 288 -----LITQCNTHTIMTGTGFIWAAYLAGGDTIYLANYPD-----SPFLKYFK 332
Db 226 DKEEAYWDMLLMQSKHGIIANSTSWRAAYL-----MEN-----PEKIIGPKHWLUF 275
Qy 333 PERAFLEFWGTIPADLSPLLKALTACPRSHFHLKAK 369
Db 276 HENICLKEWYKI-----ESHFVFKSO 296

```

## RESULT 1.4

```

US-09-479-614-14
; Sequence 14, Application US/09479614
; Publication No. US20030013183A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine
; APPLICANT: Weber, Eric
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods
; FILE REFERENCE: P-1047
; CURRENT APPLICATION NUMBER: US/09/479,614
; CURRENT FILING DATE: 2000-01-07
; EARLIER APPLICATION NUMBER: 60/115,033
; EARLIER FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 14
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Felis catus
US-09-479-614-14

```

Query Match 5.6%; Score 113; DB 9; Length 431;  
Best Local Similarity 20.6%; Pred. NO. 0.005;  
Matches 89; Conservative 40; Mismatches 127; Indels 176; Gaps 22;

```

QY 2 ASAOVPSFFPLA-----HFLIFVFWTSTIIHQOIRVKKLQPLSE 40
  ||| ||| |||
Db 4 ASIQAPLVFLPLATOCKGTIATAPSVTLGLVGYEPPVTVTWDASLNSKV-----56
  : : : : :
QY 41 KELPMITQMSG-----NTESPENRRDSEQ-----65
  ||| ||| |||
Db 57 -TLPATLOETSGLYTTTSHVTVSGEAKQKFTCSVAHAESPTINKTVSACTMNFIPTVK 115
  : : : : :
QY 66 --HNGELGRMFTINSIGRLGNQMGYATIFALARMNGRLAIFPASH-----NALAP 116
  : : : : :
Db 116 LFNS-----SCNPLGDTG-----STIQLLCLISG--YVEGDMENVTLWDGQKATN 158
  : : : : :
QY 117 IFRISLP-----VLHSDTAKKIPQNYVHLNDWMEERYRHFGHVRFTGYVPCSWTFY 168
  ||| ||| |||
Db 159 IFPYTAPKGQGVKVTSHSL-----NITQGEWVSQK-----TVTCQVTV-198
  : : : : :
QY 169 HHLRPEILKFTLHDHVREAZAQLRGLRVNGSQPSTFGVGHVHRGDYVHVMPNWKGVV 228
  ||| ||| |||
Db 199 -----QGTFEDHARKOTESDPRGVSTYLSPPSPL-----DLYVHKSPKI-TCLV 242
  : : : : :
QY 229 ADRGYLEKALDMFRARYSSPVFVTVSNGM--ANCERNINASRCDDVVFAGNLEGSPAKFE 286
  | : : : :
Db 243 VD-----LANDTGMLTWSREN-----GESVHPDP---M 268
  : : : : :
QY 287 ALLTQCNHTIMTIGTFGIWAAYLAGDITVLYANITYLPDSP--ELKVFK-PEAAFLPE-W 341
  ||| ||| |||
Db 269 VKKTQNGNITVTSTLPVDATDMWEGET-YQCKVTHFDLPKQIVRSIAKAPGRFFPEYV 327
  : : : : :
QY 342 VGIPADLSPLLK 353

```



Db 328 VFLPPEGEPTK 339

RESULT 15

US-09-479-614-2  
; Sequence 2, Application US/09479614  
; Publication No. US20030013183A1  
; GENERAL INFORMATION:  
; APPLICANT: McCall, Catherine  
; APPLICANT: Weber, Eric  
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods  
; FILE REFERENCE: P-1047  
; CURRENT APPLICATION NUMBER: US/09/479,614  
; CURRENT FILING DATE: 2000-01-07  
; EARLIER APPLICATION NUMBER: 60/115,033  
; EARLIER FILING DATE: 1999-01-07  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 496  
; TYPE: PRI  
; ORGANISM: Felis catus  
US-09-479-614-2

Query Match 5.6%; Score 113; DB 9; Length 496;  
Best Local Similarity 20.6%; Pred. No. 0.0061;  
Matches 89; Conservative 40; Mismatches 127; Indels 176; Gaps 22;

QY	2	ASAOVFFSFPLA-----HFLIFVFVTSIIHLQORIVKLQELSE	40
Db	69	ASIQAPLVFPLATCKGTATAPSVTLGCLVTGYFPMPIVTVWDARSINKSVV-----	121
QY	41	KELPMTQMSSG-----NTESEPMRRDSEQ-----	65
Db	122	-TLPATLQETSGLYTTTSHVTVSGEWAQKFTCSVAHAESPTINKTVSACTMNFIPPTVK	180
QY	66	--HGNGELRGMTINSIGRLGNQMGVATIFALARMGRLAFIPASMH-----NALAP	116
Db	181	LFHS-----SCNPLGDTG-----STQLCLISG---YVEGDMETVLVDGQKATN	223
QY	117	IFRISLP-----VLHSDTAKIPQNYHLNDWMBEYRHIFGHEVFTGYPCSWTFY	168
Db	224	IFPYTARGKQEGKVTSTHSEL-----NITQGEWWSQK-----TYICQVTY-	263
QY	169	HHLEPEILKEFTLHDHVEEAQAFLRLRVNGSQPSFVGHVHVRGDIYVHVPNNVWKGVV	226
Db	264	-----QGTEEDHARKCTESDPRGVSTYLSPPSL-----DLYVHKSPKI-TCLV	307
QY	229	ADRGYLEKALDMFRARYSSPVFVVTSGM--AWCKENINASRGDVVFAGNGIEGSPAKDF	286
Db	308	VD-----LANTGMILTWSREN-----GESVHPDP---M	333
QY	287	ALLPQCNHTIMTIGTGIWAAVLAGGDTIYLIANYILPDSF---FLKVFK-PEAAFLPE-W	341
Db	334	VKKQYNGTIVTSTPLVDADPDWVEGET-YOCKVTHTDLPKDIVRSIAKAPGRFRPPEVY	392
QY	342	VGIFADLSPLLK	353
Db	393	VFLPPEGEPTK	404

Search completed: May 27, 2003, 15:25:41  
Job time : 19.4181 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:00:45 ; Search time 13.77 Seconds  
(without alignments)  
2652.940 Million cell updates/sec

Title: US-10-040-863-8

Perfect score: 2032

Sequence: 1 MASAQVPFSPFLAHFLIFVF.....RSFHLKAKGYCYVAGRAF 380

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1485.5	73.1	343	2 A56098	alpha(1,2)fucosylt
2	1319.5	64.9	354	2 B56392	beta-galactoside a
3	1086.5	53.5	373	2 A56392	beta-galactoside a
4	1057.5	52.0	365	2 A36047	galactoside 2-alph
5	846	41.6	159	2 S46494	galactoside 2-alph
6	581	28.6	142	2 S51582	galactoside 2-alph
7	214.5	10.6	281	2 T44328	hypothetical prote
8	183	9.0	443	2 H87911	protein B0205.4 [i
9	170.5	8.4	363	2 T20745	hypothetical prote
10	156.5	7.7	348	2 T31916	hypothetical prote
11	153.5	7.6	335	2 T32294	hypothetical prote
12	146	7.2	383	2 T20572	hypothetical prote
13	136.5	6.8	500	2 T22068	hypothetical prote
14	133.5	6.6	299	2 H71976	probable alpha(1,2
15	131	6.4	365	2 T25309	hypothetical prote
16	130	6.4	353	2 T25390	hypothetical prote
17	127	6.2	388	2 T32307	hypothetical prote
18	127	6.2	392	2 T25334	hypothetical prote
19	120	5.9	625	2 T21051	hypothetical prote
20	119	5.9	434	2 T26275	hypothetical prote
21	118.5	5.8	381	2 T15140	hypothetical prote
22	117.5	5.8	365	2 T33253	hypothetical prote
23	111	5.5	395	2 T32309	hypothetical prote
24	110	5.4	371	2 T32692	hypothetical prote
25	109	5.4	277	2 T25307	hypothetical prote
26	100	4.9	317	2 T33887	hypothetical prote
27	96.5	4.7	355	2 T34405	hypothetical prote
28	93.5	4.6	348	2 T02798	hypothetical prote
29	93	4.6	597	2 G87258	acyl-CoA dehydroge

30	92.5	4.6	398	2 T19293	hypothetical prote
31	92.5	4.6	715	2 T04452	transforming prote
32	92.5	4.6	776	2 E85384	probable myb-prote
33	91	4.5	442	1 VGEEEA	glycoprotein D pre
34	89	4.4	353	2 JN0920	fibrillin precurs
35	89	4.4	468	2 S61964	probable membrane
36	89	4.4	1329	2 AB2934	peptide synthetase
37	89	4.4	1344	2 E98348	hypothetical prote
38	89	4.4	2058	2 A59267	myosin X - human
39	88.5	4.4	452	1 VGEEG3	glycoprotein D pre
40	88	4.3	217	2 A54688	modular adaptor Gr
41	88	4.3	350	2 T45082	acyl carrier prote
42	87.5	4.3	365	2 T08679	procollagen-lysine
43	87	4.3	217	2 S26050	growth factor rece
44	87	4.3	217	2 A43321	growth factor rece
45	87	4.3	438	2 T11903	NADH dehydrogenas

ALIGNMENTS

RESULT 1

A56098

alpha(1,2)fucosyltransferase Sec2, long form - human

C:Species: Homo sapiens (man)

C>Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 21-Jul-2000

C:Accession: A56098

R:Kelly, R.J.; Rouquier, S.; Giorgi, D.; Lennon, G.G.; Lowe, J.B.

J. Biol. Chem. 270, 4640-4649, 1995

A:Title: Sequence and expression of a candidate for the human Secretor blood group al

tes with the non-Secretor phenotype.

A:Reference number: A56098; MUID:95181460; PMID:7876235

A:Accession: A56098

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-343 <KE>

A:Cross-references: GB:U17894; NID:9687618; PID:AA024453.1; PID:9687619

C:Genetics:

A:Gene: GDB:FUT2; SE

A:Cross-references: GDB:120619; OMIM:182100

A:Map position: 19q13.3-19q13.3

C:Keywords: glycoprotein; Golgi apparatus; transmembrane protein

Query Match 73.1%; Score 1485.5; DB 2; Length 343;  
Best Local Similarity 77.9%; Pred.No.2.5e-116;  
Matches 275; Conservative 30; Mismatches 37; Indels 11; Gaps 2;

QY 1 MASAQVPFSPFLAHFLIFVFVSTIIHLQQRIVKLOPLSEKELPMTTOMSSGNTESEPME 60

DB 1 MLVQVMPFSPFAHFLFVFVSTIHFVQQLAKIQAM--WELFV-----QIPVLA 45

QY 61 RSEQHNGCELKGMFTINSIGRLGNOMGEYATLFPALARMNGRLAFIPASHNALADIFPI 120

DB 50 STSKALGNQSLKGMWTAIGAIGLGNOMGEYATLYALAKMNGRPAFIPAQMHSFLADIFPI 109

QY 121 SLPLVLSHDTAKKIPWQNYHLNDWMEERYRHIPGHVFRTGYPCSWTFYHHLRPEILKEFT 180

DB 110 TLPLVLSHATASRIPWQNYHLNDWMEERYRHIPGEYVRYFTGYPCSWTFYHHLRPEILKEFT 169

QY 181 LHDHVREAAQALRGLRVNGSQSPSTFGVHVRRGDYVHYVNNVWKGWVADRGYLEKALDM 240

DB 170 LHDHVREAAQALRGLRVNGSQSPSTFGVHVRRGDYVHYVNNVWKGWVADRGYLEKALDM 229

QY 241 FRARYSSPFVVTNSGMACRENIASRGDVVFAGNGIEGSPAKDFALLTQCNHIIMTIG 300

DB 230 FRARYSSPFVVTNSGMACRENIASRGDVVFAGNGIEGSPAKDFALLTQCNHIIMTIG 289

QY 301 TFGIWAAYLAGGDTIYLYANTLPDSPFLKVPKPEAAFLPEWVGIPADLSPLLK 353

DB 290 TFGIWAAYLAGGDTIYLYANTLPDSPFLKVPKPEAAFLPEWVGIPADLSPLLK 342

RESULT 2

B56392

beta-galactoside alphas,2-fucosyltransferase II - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 05-Nov-1999

C:Accession: B56392

R:Hitoshi, S.; Kusunoki, S.; Kanazawa, I.; Tsuji, S.

J. Biol. Chem. 270, 8844-8850, 1995

A:Title: Molecular cloning and expression of two types of rabbit beta-galactoside alphas

A:Reference number: A56392; MUID:95238380; PMID:7721792

A:Accession: B56392

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-354 &lt;HIT&gt;

A:Cross-references: GB:X80225; NID:g854356; PIDN:CAA56512.1; PID:g854357

C:Keywords: transmembrane protein

Query Match 64.9%; Score 1319.5; DB 2; Length 354;

Best Local Similarity 73.2%; Pred. No. 1.9e-102;

Matches 248; Conservative 27; Mismatches 47; Indels 17; Gaps 2;

QY 14 HELIFVFTSTIIHLQQRIVKQLPSEKELPMTQMSSGNTSPESMRDSEOHGNGELRG 73

DB 29 YELTFIVFVSTVFHCHQRLAL-----VPAPWAYSARVVVP-----GHLPRG 71

QY 74 MPTINSIGRLGNQMGAYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKI 133

DB 72 MWTINAMGRIGNQMGAYATLYALAKENGPRAPVIPAQMSTLAPIFRISLPVLHSDTAKKI 131

QY 134 PQWNYHLNDWMEERYHPIGHFVRTPGSCWTFYHHLRPELTKFTLHDHVRERQAQFL 193

DB 132 PQWNYHLNDWMEERYHPIGHFVRTPGSCWTFYHHLRPELTKFTLHDHVRERQAQFL 191

QY 194 RGLRYNGSPSTFFGVGHVRRGDYVHVHVNKGVVADRGYLEKALDMFRARYSSPVFVVT 253

DB 192 RGLRYNGSPSTFFGVGHVRRGDYVHVHVNKGVVADRGYLEKALDMFRARYSSPVFVVT 251

QY 254 SNGMAWCENINASRGDVVYFAGNGTEGSPAKDFALLTQCNHITMTIGTGFVWAAVLGGD 313

DB 252 SNGMAWCENINASRGDVVYFAGNGTEGSPAKDFALLTQCNHITMTIGTGFVWAAVLGGD 311

QY 314 TYILANYTLPSDPSFLKFKPEAFPEWGVIPADLSPLL 352

DB 312 TYILANYTAPDSPFLVLPKPEAFPEWGVITANMGRL 350

RESULT 3

A56392

beta-galactoside alphas,2-fucosyltransferase I - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 05-Nov-1999

C:Accession: A56392

R:Hitoshi, S.; Kusunoki, S.; Kanazawa, I.; Tsuji, S.

J. Biol. Chem. 270, 8844-8850, 1995

A:Title: Molecular cloning and expression of two types of rabbit beta-galactoside alphas

A:Reference number: A56392; MUID:95238380; PMID:7721792

A:Accession: A56392

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-373 &lt;HIT&gt;

A:Cross-references: GB:X80226; NID:g854354; PIDN:CAA56513.1; PID:g854355

C:Keywords: transmembrane protein

Query Match 53.5%; Score 1086.5; DB 2; Length 373;

Best Local Similarity 57.6%; Pred. No. 5.7e-83;

Matches 209; Conservative 51; Mismatches 62; Indels 41; Gaps 6;

QY 16 LIFVFTV-----STIIHQQRIVK-----LQPLSKKE-----LPMTTQMS----- 50

DB 11 LAFLVLCALSAFSLHLHQDLFRNGLALSPLCLERQVPAPVAIVCLPVTSPASNASSC 70

QY 51 SGNTSPESMRDSEOHGNGELRGMTINSIGRLGNQMGAYATLALARMNGRLAFIPASM 110

DB 71 AGRPAAPS-----GIWTHPDGRFGNMGQYATLLALQNLNGRRRAFLPAM 116

QY 111 HNALAPIFRISLPVLESOTAKKIPWNYHLNDWMEERYHPIGHFVRTPGSCWTFYHHL 170

DB 117 HAALAPVFRITIPVLAPEVNRRTSKQLLLHDWMBEYSRLDEPFLKFTGFCPSWTFYHHL 176

QY 171 LRPEILKFTLHDHVRERQAQFLRGV--NGSQSTFVGVRGDYVHVHVNKGVV 228

DB 177 VREQIRREFTLHDHVRERQAQFLRGV--NGSQSTFVGVRGDYVHVHVNKGVV 236

QY 229 ADRGYLERKALDMFRARYSSPVFVTSNGMAWCENINASRGDVVYFAGNGIEGSPAKDFAL 288

DB 237 GDRAYLQQAAMDWRARHEAPFVVTSGMKWCWENIDASRGDVVYFAGNGIEGSPAKDFAL 296

QY 289 LFOCNHTMTIGTGFVWAAVLGGDTIYLANFTLPDSEFLKIFKPEAFPEWGVINADL 348

DB 297 LFOCNHTMTIGTGFVWAAVLGGDTIYLANFTLPDSEFLKIFKPEAFPEWGVINADL 356

QY 349 SPL 351

DB 357 SPV 359

RESULT 4

A36047

galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) - human

C:Species: Homo sapiens (man)

C:Date: 16-Nov-1990 #sequence\_revision 16-Nov-1990 #text\_change 10-Sep-1997

C:Accession: A36047

R:Larsen, R.D.; Ernst, L.K.; Nair, R.P.; Lowe, J.B.

Proc. Natl. Acad. Sci. U.S.A. 87, 6674-6678, 1990

A:Title: Molecular cloning, sequence, and expression of a human GDP-L-fucose:beta-D-g

A:Reference number: A36047; MUID:90370848; PMID:2118655

A:Accession: A36047

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-365 &lt;LAR&gt;

A:Cross-references: GB:M35531; NID:g183887; PID:g306830

C:Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein

Query Match 52.0%; Score 1057.5; DB 2; Length 365;

Best Local Similarity 56.5%; Pred. No. 1.5e-80;

Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps 4;

QY 12 LAHFLIFVFTSTIIHLQ-----RIVKLQPLSEKELPMTQMSSGNTS 56

DB 11 LAFLVLCVLSVIFLHIHQDSPHGLGSLICPDRLVTPPVAFCLCPGTAMGNASSC 70

QY 57 PEMRDESHGNGELRGMTINSIGRLGNQMGAYATLALARMNGRLAFIPASMENALAP 116

DB 71 P-----QH-PASLSGTWTVPNGRFGNMGQYATLLALQNLNGRRRAFLPAMHAALAP 122

QY 117 IFRISLPVLHSDTAKKIPWNYHLNDWMEERYHPIGHFVRTPGSCWTFYHHLRPEIL 176

DB 123 VFRITLPVLAPEVDSTPWRELQLEDWMBEYADRLPFLKLSGFCPSWTFYHHLRPEIR 182

QY 177 KEFTLHDHVRERQAQFLRGV--NGSQSTFVGVRGDYVHVHVNKGVVADRGV 234

DB 183 REFTLHDHVRERQAQFLRGV--NGSQSTFVGVRGDYVHVHVNKGVVADRGV 242

QY 235 EKALDMFRARYSSPVFVTSNGMAWCENINASRGDVVYFAGNGIEGSPAKDFALLTQCNH 294

DB 243 EQAMDWRARHEAPFVVTSGMKWCWENIDTSOGDVTFAGDQGEATPKWOFALLTQCNH 302

QY 295 TMTIGTGFVWAAVLGGDTIYLANFTLPDSEFLKIFKPEAFPEWGVINADLSPLKA 354

DB 303 TMTIGTGFVWAAVLGGDTIYLANFTLPDSEFLKIFKPEAFPEWGVINADLSPLKA 362

QY 355 LTP 357

DB 363 AKP 365

RESULT 5

S46494  
galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) B - rat (fragment)  
N:Alternate names: alpha-1,2-fucosyltransferase  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 26-Dec-1994 #sequence\_revision 23-Feb-1996 #text\_change 08-Oct-1999  
C:Accession: S46494  
R:Piau, J.P.; Labarriere, N.; Dabouis, G.; Denis, M.G.  
Biochem. J. 300, 623-626, 1994.  
A:Title: Evidence for two distinct alpha-(1,2)-fucosyltransferase genes differentially  
A:Reference number: S46493; MUID:94280382; PMID:8010942  
A:Accession: S46494  
A:Molecule type: mRNA  
A:Residues: 1-159 <PI2>  
A:Cross-references: EMBL:L26010; NID:9414816; PIDN:AAB41515.1; PID:g414817  
C:Genetics: FTB  
A:Gene: FTB  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 41.6%; Score 846; DB 2; Length 159;  
Best Local Similarity 99.4%; Pred. No. 2.3e-63;  
Matches 158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 222 NWKGVADRGYLEKALDMFRARYSPFVVTISNGMAWCRENINASRGDVFAGNGIEGS 281  
DB 1 NWKGVADRGYLEKALDMFRARYSPFVVTISNGMAWCRENINASRGDVFAGNGIEGS 60

QY 282 PAKDFALLTOCNHTTMTIGTCIWAAYLAGDITIYIANTLPDPSFLKVFKEPAFLPEW 341  
DB 61 PAKDFALLTOCNHTTMTIGTCIWAAYLAGDITIYIANTLPDPSFLKVFKEPAFLPEW 120

QY 342 VGIPADLSPLKALTTPACPRSHFHLKAGVTCYVAGRAF 380  
DB 121 VGIPADLSPLKALTTPACPRSHFHLKAGVTCYVAGRAF 159

RESULT 6  
S51582  
galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 07-May-1995 #sequence\_revision 17-Nov-1995 #text\_change 08-Oct-1999  
C:Accession: S51582; S46493  
R:Piau, J.P.; Labarriere, N.; Dabouis, G.; Denis, M.G.  
submitted to the EMBL Data Library, November 1993  
A:Reference number: S51582  
A:Accession: S51582  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-142 <PIA>  
A:Cross-references: EMBL:L26009; NID:9414814; PIDN:AAB41514.1; PID:g554438  
Biochem. J. 300, 623-626, 1994  
A:Title: Evidence for two distinct alpha-(1,2)-fucosyltransferase genes differentially  
A:Reference number: S46493; MUID:94280382; PMID:8010942  
A:Accession: S46493  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 9-135 <PI2>  
A:Cross-references: EMBL:L26009  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 28.6%; Score 581; DB 2; Length 142;  
Best Local Similarity 73.2%; Pred. No. 2.6e-41;  
Matches 104; Conservative 15; Mismatches 21; Indels 2; Gaps 1;

QY 162 PCSWTFVHHLRPEILKEFTLHDHVREBAQAFLRGLYVNGSQPSTFVGVHVRGDIYVHV 219  
DB 1 PCSWTFVHHLRPEILKEFTLHDHVREBAQAFLRGLYVNGSQPSTFVGVHVRGDIYVHV 60

QY 220 MPNVKGVADRGYLEKALDMFRARYSPFVVTISNGMAWCRENINASRGDVFAGNGIE 279  
DB 61 MPNVKGVADRGYLEKALDMFRARYSPFVVTISNGMAWCRENINASRGDVFAGNGIE 120

QY 280 GSPAKDFALLTOCNHTTMTIGT 301

DB 121 GTFGKDFALLTOCNHTTMTIGT 142

RESULT 7  
T44328  
hypothetical protein wbla [imported] - Vibrio cholerae  
C:Species: Vibrio cholerae  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 28-Jul-2000  
C:Accession: T44328  
R:Yamasaki, S.; Shimizu, T.; Hoshino, K.; Ho, S.T.; Shimada, T.; Nair, G.B.; Takeda,  
Gene 237, 321-332, 1999  
A:Title: The genes responsible for O-antigen synthesis of Vibrio cholerae O139 are cl  
A:Reference number: Z22749; MUID:99453293; PMID:10521656  
A:Accession: T44328  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-281 <YAM>  
A:Cross-references: EMBL:AB012957; NID:g4115688; PIDN:BAA33632.1; PID:g3721682  
A:Experimental source: strain O22  
C:Genetics:  
A:Note: wbla  
C:Superfamily: Vibrio cholerae hypothetical protein wbla

Query Match 10.6%; Score 214.5; DB 2; Length 281;  
Best Local Similarity 26.5%; Pred. No. 2.5e-10;  
Matches 79; Conservative 50; Mismatches 102; Indels 67; Gaps 14;

QY 74 MFTINSIGRLNQMGAYATLFAALRMNG---RL---AFIPASMHNALAFIRISLPVLHS 127  
DB 1 MIVMKISGIGLQNFQYAVGRAIAIQYGVPLKLDYSAYKNYKLHG---YLDQFNINA 56

QY 128 DTAKKIPWQNH---NWMDEERYHI-----PGHEVREFT 159  
DB 57 DIANED--EIFHLKSSNRLSILRLRLGKNTYAEKQRTIYDVSVFMOAPRY---LD 111

QY 160 GYPCSWTFYHHLRPEILKEFTLHDHVREBAQAFLRGLYVNGSQPSTFVGVHVRGDIYVHV 219  
DB 112 GYQWNEQYFSQIRAVLLQELWPNQPLSINAGA--HQIKI---QOTHAVSIHVRGDIYVHV 165

QY 220 MPNVKGVADRGYLEKALDMFRARYSPFVVTISNGMAWCRENINASRGDVFAGNGIE 279  
DB 166 HPEIG---VLDIDYKRAVDYIKETIEAPVFVFSNDVAVCKDNENFIDSPVFIETQTE 222

QY 280 GSPAKDFALLTOCNHTTMTIGTFGIWAAYL-AGGDTIYIANTT-----LPDS 325  
DB 223 ---IDLEMLMCCOCHNIIVANSFSWAAWLNVDKIIVAPKTWMAENPKYKWPDS 277

RESULT 8  
H87911  
Protein B0205.4 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C:Accession: H87911  
R:anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C-  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;  
A:Accession: H87911  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-443 <SFO>  
A:Cross-references: GB:chr.I; PIDN:AAC16988.1; PID:g3150470; GSPDB:GN000019; CESP:B020  
A:Note: contains weak similarity to fucosyltransferases  
C:Genetics:  
A:Gene: B0205.4  
A:Map position: 1

Query Match 9.0%; Score 183; DB 2; Length 443;  
Best Local Similarity 20.5%; Pred. No. 1.9e-07;

```
Matches 72; Conservative 53; Mismatches 128; Indels 96; Gaps 11;
QY 83 LGNOMGEYATLAFALARMGRIFAPA--SMHNA--LAFIFRISLPVLHSDPAKTI----- 133
Db 103 IGLNMFQVAGLLSIARETGSIILLISTTILRAPDFEITFNDISQFVGEDLSRQLAEDLN 162
QY 134 -----PWQNYHLNDWMEERYRHIPGHVRTGTPCSTWTFYHHLRPEILKE-FTLH 182
Db 163 ASKITLTSCCAYNRLSTILFNDISRI-----IERIDGYFQNFYRPHPSQKIVKLFTEM 216
QY 183 DHVREAAQAF-----LRGLR 197
Db 217 DPVRKRVFNNLIYNIHPTNHRKPKSTVSIISFVFPQLRVDLENIGLSLTVRNAR 276
QY 198 V-----NGSQ-----PSTFGVHVHRRGDYHVH--MPNVWKGWVADRGYLEKAL 238
Db 277 VIETNVANDQALELPEEDAFAKTMWGVHIRHGMDSMNSERNRIGHVDPIEYKKRAI 336
QY 239 DMFRARYSSPVVTSNGMAWRENINASRGDVPFAGNGIGSFAKDPALLTQCCHTMT 298
Db 337 QOISKIYENVAFIICSDNVAVARRNMLKGLKTLHFF---CPGPREVDMAILKSCDSVIIS 393
QY 299 IGTFQIWAAYL---AGGDTIYLANVTLDPSPFLKVFKEPAFLPEWGI 344
Db 394 TGTGWSAYLNVASPDVYIYKHPAPGWSMERKNTKTEYFLKSWTAL 442

RESULT 9
T20745
hypothetical protein F11A5.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T20745
R:Gardner, A.
submitted to the EMBL Data Library, March 1997
A:Reference number: Z19319
A:Accession: T20745
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-363 <WIL>
A:Cross-references: EMBL:Z292830; PIDN:CA07352.1; GSPDB:GN00023; CESP:F11A5.5
A:Experimental source: clone F11A5
C:Genetics:
A:Gene: CESP:F11A5.5
A:Map position: 5
A:Introns: 42/1; 55/2; 136/3; 175/3; 210/2; 243/1; 299/1; 330/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 8.4%; Score 170.5; DB 2; Length 363;
Best Local Similarity 23.4%; Pred. No. 1.7e-06;
Matches 73; Conservative 45; Mismatches 101; Indels 93; Gaps 17;
QY 79 SIGRLNOMGEYATLAFALARMGRILA--FIPASMH-----NALAFIFRISLPVLHSDT 129
Db 78 SSRGLNHLFELASVLSIRELQRPVPTFEIENCYHEKMWEDSNTLIPGLNHLPLINGSV 137
QY 130 AKKIPWQNYH-----LNDWMEERYRHIPGHVRTGTPCSTWTFYHHLRPEILKEF 179
Db 138 PSSVKRVYFKQKCTDDPSLLDNYDEYLLHLAG-----THYQ-SMKYFSHMRNELI--- 188
QY 180 ILHDHVEEAQAFLGLRVNGSQPSTFVG-VHVRRGDYHVHMPNVWKGW-VADRGYLEKA 237
Db 189 ---GYLKTENTY---MDLPKSGENTFITCVHVRGDFLRV-----GFHVADENFIRSS 236
QY 238 LDMFRARYSSPVFWVTSNGMAWRENINASRGDVPFAGNGIE-----VFGDDYEDWLSLRNRTSKINAFVSQ 278
Db 237 LNLISQVAKRANTAT-----VFGDDYEDWLSLRNRTSKINAFVSQ 278
QY 280 GSPAKDFALL-TCQNTHTMTI--GTGFIWAAYLAGDGTIYLANVTLDPSPFLK----- 329
Db 279 NSPADLLLYAKSCNDVVLTAHAHSTFCGWMGYFSKGNRVY--YT--DIQFTKDWILETG 333
QY 330 VFKEPAFLPEW 341
```

```
Db 334 EFISEDYILPHW 345
RESULT 10
T31916
hypothetical protein C17A2.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jul-2000
C:Accession: T31916
R:Sammons, L.; Wohlmann, P.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid C17A2.
A:Reference number: Z21098
A:Accession: T31916
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-348 <SAM>
A:Cross-references: EMBL:AF016654; PIDN:AAB66030.1; GSPDB:GN00023; CESP:C17A2.4
A:Experimental source: strain Bristol N2; clone C17A2
C:Genetics:
A:Gene: CESP:C17A2.4
A:Map position: 2
A:Introns: 94/3; 133/3; 168/2; 272/1
C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 7.7%; Score 156.5; DB 2; Length 348;
Best Local Similarity 20.9%; Pred. No. 2.3e-05;
Matches 72; Conservative 57; Mismatches 118; Indels 97; Gaps 18;
QY 12 LAHFLIF-VFVTSTIIHLQQRIVKLQPLSEKELPMTTQSSNGTESPEMRDSEQHNGE 70
Db 6 VVHSMERFVPPAVLHENS-----TEHFLSNLASP----- 37
QY 71 LRGMFTINSIGRLNOMGEYATLAFALARMGR--LAFIPASMHNALAFIFRISLP----- 123
Db 38 -----SRGLNHIFELASFLGLSERLHRTPLFLVENEHFQKMLDTRKVPGLVEK 87
QY 124 --VLHSDTAKI---PWO-----NYHLNDWMEERYRHIPGHVRTGTPCSTWTFYHHL 171
Db 88 FTVINGSLSIFSIKTPQKVCCHRENFEILEKIDKYLHLTGMFIQ-----SNKYFENM 141
QY 172 RPEILKEFTLHDHVEEAQAFLGLRVNGSQPSTFVG-VHVRRGDYHVHMPNVWKGW-VAD 230
Db 142 REQLL-----DFLDDSSQDF--CNLPRSQRTHVTCVHARRGDFVDV-----GFOAAD 187
QY 231 RGYLEKALDMFRARYSSPV-----FVYTSNGMAWCR---EN--INASRGDVPFAGNG 277
Db 188 PDFIRNSVKYIAENFIPELEYKVKYHRKVIFGDDLEFMRSLFNSVSTDEPEYMFPAEY 247
QY 278 I--EGSPAKDFALLTQ--CNHTIMTI--GTGFIWAAYLAGDGTIY 316
Db 248 YISQNSPAEDLTYSKQCDIVLISAPASTFGWIGYFSKGNKVF 291

RESULT 11
T32294
hypothetical protein K06H6.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
C:Accession: T32294
R:David, M.; Wohlmann, P.; Bauer, C.; Clarke, K.
submitted to the EMBL Data Library, September 1997
A:Description: The sequence of C. elegans cosmid K06H6.
A:Reference number: Z21147
A:Accession: T32294
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-335 <DAV>
A:Cross-references: EMBL:AF024500; PIDN:AAB70365.1; GSPDB:GN00023; CESP:K06H6.6
A:Experimental source: strain Bristol N2; clone K06H6.6
C:Genetics:
A:Gene: CESP:K06H6.6
```

RESULT 14  
H71976  
probable alpha(1,2)fucosyltransferase - Helicobacter pylori (strain J99)  
C:Species: Helicobacter pylori  
A:Variety: strain J99  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 28-Jul-2000  
C:Accession: H71976  
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Deig, P.C.;  
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.;  
Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
A:Reference number: A71800; MUID:99120557; PMID:9923682  
A:Accession: H71976  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-299 <ARN>  
A:Cross-references: GB:AB001447; GB:AB001439; NID:94154583; PIDN:AA005659.1; PID:9415458  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: jhp0086  
C:Superfamily: Vibrio cholerae hypothetical protein wbia

Query Match 6.4%; Score 133.5; DB 2; Length 299;  
Best Local Similarity 22.9%; Pred. No. 0.0016;  
Matches 76; Conservative 38; Mismatches 129; Indels 89; Gaps 13;

```
QY 81 GRLGNMGGEYA-----TFLARANGRLAFIPASMHNALAPIFRIS-- 121
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 10 GGLGNMGQYAFKSLKHSNTPVLLDITSPGSNKKMOLEFPIDLPVSAKEIAIAKM 69
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 122 --LPVHSDTAKKIPQNNVHLNDWMEERYRHI-PGHFVFTGYPCSWTFYHHLRPEILKE 178
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 70 QHLPKLVDAKLYMGFDVRSQEIIVEYEPKLLKPSRLATYFYGFQDPFYDAISSLIKQT 129
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 179 FTL-----HDHVREEAQAFRLGLR-VNGSQPSFTFVGHVHVRGDIYVHVMNVKGVVADRG 232
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 130 FTLPFPPENGNNKKEEYHKKLSILAKNSVF--AHIRRGDIYGI-----GCQLGID 181
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 233 YLEKALDMFRARYSSPVFVVTNGMAWCRENINASRGDVVFAGNGIEGSPAKDFA----- 287
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 182 YQKKAVEYMAKRVPNMELFV-----FCB-----DLKFTQNLDLGYPFMDMTTRDKD 227
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 288 -----LLTCNHTIMTIGTGWAAAYLAG--GCTIYLANITLPDSPFLKVFPEPAFAF 337
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 228 EYAYDMLLMQSCCKGIIANSYSYWAAYLINNPKKII-----GPKHWFGLGHENIL 279
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 338 LPEWVGIPADLSPLLKALTACPRSHFLKAK 369
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 280 CKEWKI-----ESHFEVKSQ 295
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 15  
T25309  
hypothetical protein T26E4.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jul-2000  
C:Accession: T25309  
R:McMurray, A.  
submitted to the EMBL Data Library, October 1996

A:Reference number: Z20013  
A:Accession: T25309  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-365 <WIL>  
A:Cross-references: EMBL:Z81132; PIDN:CAB03434.1; GSPDB:GN00023; CESP:T26E4.4  
A:Experimental source: clone T26E4  
C:Genetics:  
A:Gene: CESP:T26E4.4  
A:Map position: 5  
A:Introns: 50/1; 80/2; 140/3; 179/3; 218/2; 253/1; 278/1; 306/3; 325/1; 342/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 6.4%; Score 131; DB 2; Length 365;  
Best Local Similarity 20.7%; Pred. No. 0.0033;  
Matches 61; Conservative 43; Mismatches 106; Indels 84; Gaps 13;

```
QY 84 GNGMGVATLFLARANGRLAFIPA-----SMHNALAPIFRISLPVLSHSDTAKKIPWQ 136
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 86 GNKLFEILISLLGIANSLQRPVIDAIPSNIRSLHKSIGLFP---KLVEQFDLKMIPAS 142
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 137 NY--HLNDWME-----EYRHIPGHVFTGYPCSWTFYHHLRPEILKEFT 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 143 SVSSHQNNWVKCCIFDDPKMLNRSQHLMLNGHYTFQ-----SFKYFHLHURSEIREWLA 196
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
QY 181 LHDHVREEAQAFRLGLRVNGSQPSFTFVGHVHVRGDIYVHVMNVKGVVADRGYLEK 236
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 197 PSKMAKLAETVL-----TSELKEDLIICTHIRRGDFQIDGVH-QP-----SDPNFTRA 244
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 237 ALDMFRARYSS-----PVFVVTSGMAWCRENIN-----NASR 268
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 245 ATDFLVKKHYOKHHYETTVVVTGNDYNFASKAFEDRVSNSSVIPNRTTPPLNFPPIENSEPK 304
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 269 GDVVFAGNGIEGSPAKDFAALLTCNHTIMTIGTGWAAAYLAGGDTIYLANITL 322
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 305 YSVILLPON---STPENDLAFSRQAPSS-----TFGWLSYLAKRSAYILRPFLCL 350
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: May 27, 2003, 15:08:16  
Job time : 28.77 secs

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:53:11 ; Search time 11.122 Seconds  
(without alignments)  
1417.108 Million cell updates/sec

Title: US-10-040-863-8

Perfect score: 2032

Sequence: 1 MASQVFFSEPLAHFLIFV.....RSHFLKAKGVTCYVAGRAF 380

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1502	73.9	340	1	FUT2_PIG
2	1485.5	73.1	343	1	FUT2_HUMAN
3	1466	72.1	344	1	FUT2_BOVIN
4	1363.5	67.1	368	1	FUT2_MOUSE
5	1319.5	64.9	354	1	FUT2_RABIT
6	1107	54.5	376	1	FUT1_MOUSE
7	1102	54.2	376	1	FUT1_RAT
8	1086.5	53.5	373	1	FUT1_RABIT
9	1057.5	52.0	365	1	FUT1_HUMAN
10	1050.5	51.7	365	1	FUT1_PIG
11	846	41.6	159	1	FUT2_RAT
12	168	8.3	365	1	YKQ7_CAEEL
13	91	4.5	442	1	VGLD_HSVK
14	89	4.4	353	1	FMA3_PORGI
15	89	4.4	773	1	CDH_PHACH
16	89	4.4	2058	1	MY10_HUMAN
17	88.5	4.4	359	1	PTN7_RAT
18	88.5	4.4	402	1	VGLD_HVEA
19	88.5	4.4	452	1	VGLD_HSVR
20	88	4.3	217	1	GRB2_MOUSE
21	87	4.3	217	1	GRB2_HUMAN
22	87	4.3	438	1	NQOL_THERH
23	87	4.3	533	1	QUPD_EKENI
24	86.5	4.3	1112	1	CN3B_HUMAN
25	85.5	4.2	738	1	PL03_HUMAN
26	85.5	4.2	1928	1	LPH_RAT
27	84.5	4.2	344	1	RIBB_HELPY
28	84.5	4.2	383	1	ATP_THAQ
29	84.5	4.2	718	1	CTPC_MYCTU
30	84	4.1	739	1	AB1_MYCTU
31	84	4.1	1780	1	POLG_MVEV
32	83.5	4.1	273	1	TC1A_CAEEL
33	83.5	4.1	350	1	FME2_PORGI

## RESULT 1

ID	FUT2_PIG	STANDARD	PRT	340 AA.
AC	Q10982: Q29044; O19100;			
DT	01-OCT-1996 (Rel. 34; Created)			
DT	15-JUL-1999 (Rel. 38; Last sequence update)			
DT	15-JUN-2002 (Rel. 41; Last annotation update)			
DE	Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)FT 2)			
DE	(Fucosyltransferase 2).			
GN	FUT2.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97468270; PubMed=9321466;			
RA	Meijerink E., Fries R., Voegeli P., Masabanda J., Wigger G., Stricker C., Neuenchwander S., Bertschinger H.U., Stranzinger G.;			
RT	"Two alpha(1,2) fucosyltransferase genes on porcine chromosome 6q11 are closely linked to the blood group inhibitor (S) and Escherichia coli F18 receptor (ECF18R) loci.";			
RT	Mamm. Genome 8:736-741(1997).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	Cohney S., Mouhtouris E., McKenzie I.F.C., Sandrin M.S.;			
RA	"Molecular cloning and characterization of the pig secretor type alpha(1,2)fucosyltransferase.";			
RT	Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.			
RL	[3]			
RP	SEQUENCE OF 69-334 FROM N.A.			
RA	Petit J.M.;			
RL	Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE OF 69-80; 119-133 AND 316-334.			
RC	TISSUE=Submaxillary gland;			
RX	MEDLINE=96064742; PubMed=7592879;			
RA	Thurin J., Blaszczyk-Thurin M.;			
RT	"Porcine submaxillary gland GDP-L-fucose: beta-D-galactoside alpha-2-L-fucosyltransferase is likely a counterpart of the human secretor gene-encoded blood group transferase.";			
RT	J. Biol. Chem. 270:26577-26580(1995).			
RL	CC			
CC	-!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP + alpha-L-fucosyl-1,2-beta-D-galactosyl-R.			
CC	-!- PATHWAY: Glycosylation.			
CC	-!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).			

P05418 paracoccus  
Q61409 mus musculus  
P76464 escherichia  
P45784 vibrio chol  
P72933 synechocyst  
P87503 human adeno  
P97846 rattus norv  
P78357 homo sapien  
P87379 xenopus lae  
Q92140 helicobacte  
P14014 bacillus li  
P23739 rattus norv

CYB\_PARDE  
CN3B\_MOUSE  
VFAS\_ECOLI  
GSPN\_VIBCH  
GLO2\_SYNY3  
DPO2\_ADE04  
CTAL\_RAT  
CTAL\_HUMAN  
GRB2\_XENLA  
RIBB\_HELPY  
CDGT\_BACLI  
SUIS\_RAT

1 440  
1 799  
1 1534  
1 252  
1 257  
1 1193  
1 1381  
1 1384  
1 217  
1 344  
1 718  
1 1840

83.5 4.1  
83.5 4.1  
83.5 4.1  
83 4.1  
82.5 4.1  
82.5 4.1  
82.5 4.1  
82.5 4.1  
82 4.0  
82 4.0  
82 4.0  
82 4.0  
82 4.0

## ALIGNMENTS



```

CC EMBL: U70881; AAB81883.1; -
DR EMBL: AF027304; AAC09170.1; -
DR EMBL: X99621; CAA67932.1; -
DR InterPro: IPR002516; GT_11
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Golgi stack; Signal-anchor.
FT DOMAIN 1 7 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 8 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT
FT DOMAIN 29 340 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 279 279 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 305 305 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 158 158 R -> H (IN REF. 3).
SQ SEQUENCE 340 AA; 38987 MW; 0629FIC04FC206AD CRC64;

Query Match 73.9%; Score 1502; DB 1; Length 340;
Best Local Similarity 79.6%; Pred. No. 1.3e-119;
Matches 281; Conservative 28; Mismatches 30; Indels 14; Gaps 3;

QY 1 MASAOVFSPFLAHLFLFVFTIHLQRIYKLOPLSEKELPMTQSSGNTSEPMR 60
Db 1 MLSQASFFPTGPFILFVFTASTIFHLQRMVKIQP--TWELQMTQVT---TESP--- 52
QY 61 RDSQHGNGELRGMTINSIGRLGNQMGVATLALARMNGRAFLPASMNALAFIRI 120
Db 53 -----SSPOLKGMWITNALIGRLGNQMGVATLYALARNGRPAFLPPMHSHTLAFIRI 106
QY 121 SLPLVLSHDTAKKIPWQNYHLNDWMEERYRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFT 180
Db 107 TDLVLHASTARRIPWQNYHLNDWMEERYRHIPGEYVRLTGYPSCSWTFYHHLRPEILKEFT 166
QY 181 LHDHVRERACAFURGLVNGSOPSTGVVVRGDDVHVHVPNNVKGVDRCGLVKALDM 240
Db 167 LNHVREERQDFLRLVNGSRPSTVGVVRGDDVHVHVPNNVKGVDRCGLVKALDM 226
QY 241 FRARYSPVFTVSSNGMARENINASRGDVFVAGNIEGSPAKDFALLTQCNTHTMTIG 300
Db 227 FRARYSPVFTVSSNGMARENINASRGDVFVAGNIEGSPAKDFALLTQCNTHTMTIG 286
QY 301 TFCIMAAVLGGDTIYLANITLPDSFLKVPKPAFLPFWGVPADLSPLLK 353
Db 287 TFCIMAAVLGGDTIYLANITLPDSFLKVPKPAFLPFWGVPADLSPLLK 339

RESULT 2
FUT2_HUMAN
ID FUT2_HUMAN STANDARD; PRT; 343 AA.
AC Q10981;
DT 01-OCT-1996 (Rel. 34; Created)
DT 01-OCT-1996 (Rel. 34; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (GDP-L-fucose:beta-
DE D-galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)Ft 2)
DE (Fucosyltransferase 2) (Secretor blood group alpha-2-
DE fucosyltransferase) (Secretor factor) (Se) (SE2).
GN FUT2 OR SEC2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human Secretor blood
RT group alpha(1,2)fucosyltransferase gene (FUT2). Homozygosity for an
RT enzyme-inactivating nonsense mutation commonly correlates with the
RT non-secretor phenotype."
RL J. Biol. Chem. 270:4640-4649(1995).
```

```

RN [2]
RP SEQUENCE FROM N.A., AND VARIANT PHE-140.
RX MEDLINE=96199252; PubMed=8621666;
RA Kudo T., Iwasaki H., Nishihara S., Shinya N., Ando T., Narimatsu I.,
RA Narimatsu H.;
RT "Molecular genetic analysis of the human Lewis histo-blood group
RT system. II. Secretor gene inactivation by a novel single missense
RT mutation A395T in Japanese nonsecretor individuals."
RL J. Biol. Chem. 271:9830-9837(1996).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT PHE-140.
RX MEDLINE=97363210; PubMed=9219535;
RA Koda Y., Soejima M., Wang B., Kimura H.;
RT "Structure and expression of the gene encoding secretor-type
RT galactoside 2-alpha-L-fucosyltransferase (FUT2).";
RL Eur. J. Biochem. 246:750-755(1997).
RN [4]
RP VARIANTS VAL-25; CYS-138 AND ASN-172.
RX MEDLINE=98431007; PubMed=9760207;
RA Liu Y., Koda Y., Soejima M., Pang H., Schlaphoff T., du Toit E.D.,
RA Kimura H.;
RT "Extensive polymorphism of the FUT2 gene in an African (Xhosa)
RT population of South Africa."
RL Hum. Genet. 103:204-210(1998).
CC !- FUNCTION: CREATES A SOLUBLE PRECURSOR OLIGOSACCHARIDE FUC-ALPHA
CC ((1,2)GALBETA-) CALLED THE H ANTIGEN WHICH IS AN ESSENTIAL
CC SUBSTRATE FOR THE FINAL STEP IN THE SOLUBLE A AND B ANTIGEN
CC SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE THE SAME ACCEPTOR
CC SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.
CC !- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC !- PATHWAY: Glycosylation.
CC !- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC !- ALTERNATIVE PRODUCTS: TWO FORMS ARE PRODUCED BY THE USE OF
CC ALTERNATIVE INITIATION CODONS IN THE SAME READING FRAME.
CC !- TISSUE SPECIFICITY: SMALL INTESTINE, COLON AND LUNG.
CC !- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
CC TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF
CC MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.
CC !- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL: U17894; AAC24453.1; -
DR EMBL: D89327; BAA13944.1; -
DR EMBL: D87942; BAA21684.1; -
DR MIM: 182100; -
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack; Alternative initiation;
KW Blood group antigen; Polymorphism.
FT CHAIN 1 343 GALACTOSIDE 2-L-FUCOSYLTRANSFERASE 2,
FT LONG ISOFORM.
FT CHAIN 12 343 GALACTOSIDE 2-L-FUCOSYLTRANSFERASE 2,
FT SHORT ISOFORM
FT INIT_MET 12 12 FOR SHORT ISOFORM.
FT DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 15 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 29 343 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 282 282 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 25 25 I -> V (IN XHOSA POPULATION).
FT /FTID=VAR_003422.
```

```

FT VARIANT 138 138 R -> C (IN XHOSA POPULATION).
FT FTID=VAR_003423.
FT VARIANT 140 140 I -> F (IN JAPANESE SEJ ALLELE; NON-
FT SECRETOR).
FT FTID=VAR_003424.
FT VARIANT 172 172 D -> N (IN XHOSA POPULATION).
FT FTID=VAR_003425.
SQ SEQUENCE 343 AA; 39017 MW; 12066D9CF175E13A CRC64;

Query Match 73.1%; Score 1485.5; DB 1; Length 343;
Best Local Similarity 77.9%; Pred. No. 3.4e-118;
Matches 275; Conservative 30; Mismatches 37; Indels 11; Gaps 2;

QY 1 MASQVPSFPLAHFLIFVFTSTIIHLQQRIVKLQPLSEKELPMTQMSSGNTESPEMR 60
DB 1 MLVQMPFSPMAHFILEVFTVSTIFVHQRLAKIQAM--WELPV-----QIPVLA 49

QY 61 RDSQHGNGELRGMTINSTGRGNQMGEXATLFAALRMNGRLAFIPASMHNALAPIFRI 120
DB 50 STKALGPSOLRGMTTINAIIGRLGNQMGEXATLYALAKMNGRPAFIPAQMHSHTLAPIFRI 109

QY 121 SLPLVHSDTAKKIPWQNYHLNDWMEERYRHIPGHFVFTGYPCSWTFYHHLRPEILKEFT 180
DB 110 TLPVLHSATSRIPWQNYHLNDWMEERYRHIPGHFVFTGYPCSWTFYHHLRPEILKEFT 169

QY 181 LHDHVRERAAQFLRGLRVNGSQPSTFYGVHVRGDDYVHVMPNVKGVVADRGYLEKALDM 240
DB 170 LHDHVRERAAQFLRGLRVNGSRPCTFYGVHVRGDDYVHVMPNVKGVVADRGYLEKALDM 229

QY 241 FRARYSPFVFTVNSGMWACREINASRGDVFVAGNIEGSPAKDFALLTCCNHTIMTIG 300
DB 230 FRARYSLIFVFTVNSGMWACREINIDTSHGDVFGAGDIEGSPAKDFALLTCCNHTIMTIG 289

QY 301 TFGIWAAYLAGGDTIYLYANTLPDPSPLKPKFAAFLPEWVGIPADLSPLLK 353
DB 290 TFGIWAAYLAGGDTIYLYANTLPDPSPLKPKFAAFLPEWVGIPADLSPLLK 342

RESULT 3
FUT2_BOVIN STANDARD; PRT; 344 AA.
AC Q28113;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (GDP-L-fucose:beta-
DE D-galactoside 2-alpha-L-fucosyltransferase 2) (alpha(1,2)FT 2)
DE (Fucosyltransferase 2).
GN FUT2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Petit J.M.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X99620; CAA67931.1; -.

```

```

DR InterPro: IPR002516; GI 11.
KW Pfam: PF01531; Glyco_transf_11; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Golgi stack; Signal-anchor.
FT DOMAIN 1 7 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 8 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 29 344 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 283 283 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 344 AA; 39320 MW; F35AC33F1B7B9F68 CRC64;

Query Match 72.1%; Score 1466; DB 1; Length 344;
Best Local Similarity 77.1%; Pred. No. 1.5e-116;
Matches 272; Conservative 27; Mismatches 44; Indels 10; Gaps 1;

QY 1 MASQVPSFPLAHFLIFVFTSTIIHLQQRIVKLQPLSEKELPMTQMSSGNTESPEMR 60
DB 1 MFSTCTFEFFPTAFILFVFTASTIFHLHQRLKMQPTWELEA-----LEPATME 50

QY 61 RDSQHGNGELRGMTINSTGRGNQMGEXATLFAALRMNGRLAFIPASMHNALAPIFRI 120
DB 51 TPRSPPRPQPLKGMWTINAIIGRLGNQMGEXATLYALAKMNGRPAFIPAQMHSHTLAPIFRI 110

QY 121 SLPLVHSDTAKKIPWQNYHLNDWMEERYRHIPGHFVFTGYPCSWTFYHHLRPEILKEFT 180
DB 111 TLPVLHSDTAKSPWPQNYHLNDWMEERYRHIPGHFVFTGYPCSWTFYHHLRPEILKEFT 170

QY 181 LHDHVRERAAQFLRGLRVNGSQPSTFYGVHVRGDDYVHVMPNVKGVVADRGYLEKALDM 240
DB 171 LHAHVRERAAQFLRGLRVNGSRPSTFYGVHVRGDDYVHVMPNVKGVVADRGYLEKALDM 230

QY 241 FRARYSPFVFTVNSGMWACREINASRGDVFVAGNIEGSPAKDFALLTCCNHTIMTIG 300
DB 231 FRARYSAPFVFTVNSGMWACREINASRGDVFVAGNIEGSPAKDFALLTCCNHTIMTIG 290

QY 301 TFGIWAAYLAGGDTIYLYANTLPDPSPLKPKFAAFLPEWVGIPADLSPLLK 353
DB 291 TFGIWAAYLAGGDTIYLYANTLPDPSPLKPKFAAFLPEWVGIPADLSPLLK 343

RESULT 4
FUT2_MOUSE STANDARD; PRT; 368 AA.
AC P97333;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (GDP-L-fucose:beta-
DE D-galactoside 2-alpha-L-fucosyltransferase 2) (alpha(1,2)FT 2)
DE (Fucosyltransferase 2) (FUT-III).
GN FUT2 OR SECL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hitoshi S.; Kusunoki S.; Kanazawa I.; Tsuji S.;
RC STRAIN=ICR;
RT "Molecular cloning and expression of a mouse GDP-L-fucose: beta-D-
RT galactoside 2-alpha-L-fucosyltransferase.";
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR;
RA Lin B.; Hayashi Y.; Saito M.; Sakakihara Y.; Yandagisawa M.;
RA Wamori M.;
RT "Molecular cloning and expression of a GDP-L-fucose:beta-D-galactoside
RT 2-alpha-L-fucosyltransferase in murine gastrointestinal tract.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

```



```
Db 252 SNGMWCNRENIDASRGVDFVAGNGLEGPSPAKDFALLTQCNTHTVMTIGTFGFWAAYLTGGD 311
QY 314 TIVIANVTLSDSPFLKVKPEAAFLPEWVGIPADLSPLL 352
Db 312 TVILANTAPDSFHLVFKPEAAFLPEWVGITANKGRAL 350
RESULT 6
ID FUT1_MOUSE STANDARD; PRT; 376 AA.
AC O09160;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.65) (GDP-L-fucose:beta-
D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)Ft 1)
DE (Fucosyltransferase 1).
GN FUT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1].
RP SEQUENCE FROM N.A.
RC STRAIN=NTH Swiss;
RX MEDLINE=97454449; PubMed=9355741;
RA Domino S.E., Hiraiwa N., Lowe J.B.;
RT "Molecular cloning, chromosomal assignment and tissue-specific
expression of a murine alpha(1,2)fucosyltransferase expressed in
thymic and epididymal epithelial cells.";
RL Biochem. J. 327:105-115(1997).
CC -1- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -1- PATHWAY: Glycosylation.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
FORM IN TRANS CISTERNAE OF GOLGI.
CC -1- TISSUE SPECIFICITY: IN THE ADULT, HIGHLY EXPRESSED IN PANCREAS,
TESTIS AND EPIDIDYMIS AND TO A LESSER EXTENT IN THYMUS, LONG,
STOMACH, SMALL INTESTINE, COLON, SPLEEN AND UTERUS. NOT EXPRESSED
IN BRAIN, HEART, SKELETAL MUSCLE, KIDNEY, LIVER AND BONE MARROW.
CC -1- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
GALACTOSIDE 2-L-FUCOSYLTRANSFERASE.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: U90553; AAC53492.1; -
DR MGD; MGI:109375; Fut1.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
Signal-anchor; Golgi stack.
FT DOMAIN 1 8 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 9 26 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 27 376 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 67 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 302 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 328 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 376 AA; 42255 MW; 21FD24CFE204106 CRC64;
Query Match 54.5%; Score 1107; DB 1; Length 376;
Best Local Similarity 60.5%; Pred. No. 3.6e-86;
Matches 210; Conservative 44; Mismatches 83; Indels 10; Gaps 3;
QY 21 VTSIIHQRIYKQLPSKELPMVTMTQSSNGTSEPMRDSQHNGELRGFTTNSI 80
Db 38 LTUSVLCSYHLIK-SPVAMVCLPPLPQTSGNSPCEQ-----SSLSGTWITTPG 88
```

```
QY 81 GRLGNMGVATLALARMNGRLAPIPASMINALAPIPRIISLPVLHSTAKKIPRONIHL 140
Db 89 GRFGNGMGQYATLLALAOQLNGRQAFIOPEMHAALAPVPRISLPVLDPEDVSLTPWQHLVL 148
QY 141 NDWMEERYRHIPGHPVFTGPGCSWTFYVHHLRPELTKFTLHDHVRBEAQAFRLGLRVNG 200
Db 149 HDWMESEYSHLSDPEFLKSGPFCSTWTFHHLREQIRREFTLHNLHRLDGAQYLLSGLRGP 208
QY 201 SQPS-IFVGVHVRGDDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAW 259
Db 209 ASPAHTFVGVHVRGDDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAW 268
QY 260 CREMINASRGDVGAVAGNIEGSPAKDFALLTQCNTHTVMTIGTFGFWAAYLAGGDIIVLAN 319
Db 269 CLENIDTSHGDDVFGAGNGQEGTGGKDFALLTQCNTHTVMTIGTFGFWAAYLAGGDIIVLAN 328
QY 320 YTLPSDFLKVFKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFL 366
Db 329 FTLPDSFLKTFRPEAAFLPEWVGINADLSPLQAFDPWKDPSLFLRL 375
```

## RESULT 7

```
ID FUT1_RAT STANDARD; PRT; 376 AA.
AC Q10960;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.65) (GDP-L-fucose:beta-
D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)Ft 1)
DE (Fucosyltransferase 1).
GN FUT1 OR FTA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1].
RP SEQUENCE FROM N.A.
RC TISSUE-Colon cancer;
RA Soejima M., Wang B., Koda Y., Kimura H.;
RT "Two distinct rat GDP-L-fucose:b-D-galactoside 2-alpha-L-
fucosyltransferase genes.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2].
RP SEQUENCE OF 169-310 FROM N.A.
RC STRAIN=BDIX;
RX MEDLINE=94280382; PubMed=8010942;
RA Piau J.-P., Labarriere N., Dabouis G., Denis M.G.;
RT "Evidence for two distinct alpha(1,2)-fucosyltransferase genes
differentially expressed throughout the rat colon.";
RL Biochem. J. 300:623-626(1994).
CC -1- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -1- PATHWAY: Glycosylation.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).
CC -1- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF
MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: AB015637; BAA31130.1; -
DR EMBL; AB006137; BAA21741.1; -
```

```

DR EMBL; L26009; AAB41514.1; -.
DR InterPro; IPR002516; GT.11.
DR Pfam; PF01531; Glyco.transf.11; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 13 29 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).
FT DOMAIN 30 376 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 64 64 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 376 AA; 42416 MW; C350C737C758B7F8 CRC64;

Query Match 54.2%; Score 1102; DB 1; Length 376;
Best Local Similarity 60.1%; Pred. No. 9.6e-86;
Matches 209; Conservative 43; Mismatches 84; Indels 12; Gaps 3;

QY 21 VSTIIHLQORIVKQLPSKELPMTQSSGNTESEPMRROSEHQHNGELRGMTFINSI 80
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 38 LTFSLVCPDXYHLK-SPVAMVCLPFPNSASSGSPSCPE-----QSLLSGTWITPG 87

QY 81 GRIGNOMGEVATIFALARNGRGLAFIPASMINALAFIRISLPVLSHSDAKKIPWQNYHL 140
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 88 GRFGNMGQYATLLAQLNGRAFRATFQPEHMTTLAPVFRISLPVLPDPEVDSLPWQHVL 147

QY 141 NDWMEERYRHIPGHVRFVTCYPCSWTFYHHRPEILKEPTLHDHVEEAQAFRLRVN- 199
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 148 HDWMSSEYSHLEDFPKLSGFPSCWTFHHLRQIRREFTLHDHREDAGRLSLGRIGP 207

QY 200 -GSPSTFGVGVHVRGDIYHVMNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMA 258
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 208 AGIRPRTYGVHVRGDIYHVMNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMR 267

QY 259 WCENINASEGDIYHVRGDIYHVMNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMA 318
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 268 WCLENIDTSHGVVAGNQGEGTGGKDFALLTQCNHTIMTIGTFGWAAYLAGDVAFLA 327

QY 319 NYTLPDPSFLKFKPAAPFLPEWVGIPADLSPLLKALTPACPRSHFL 366
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 328 NFTLPDSEFLKFKPAAPFLPEWVGIPADLSPLLKALTPACPRSHFL 375

RESULT 9
FUT1_HUMAN
ID FUT1_HUMAN STANDARD; PRT; 373 AA.
AC Q10979;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)
DE (Fucosyltransferase 1).
DE FUT1 OR RFN-1.
GN Oryctolagus cuniculus (Rabbit).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95238380; PubMed=7721792;
RA Hitoishi S., Kusumoki S., Kanazawa I., Tsuji S.;
RT "Molecular cloning and expression of two types of rabbit beta-galactoside alpha 1,2-fucosyltransferase.";
RL J. Biol. Chem. 270:8844-8850(1995).
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP + alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC -!- FORM IN TRANS CISTERNAE OF GOLGI.
CC -!- TISSUE SPECIFICITY: BRAIN.
CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE

```

```

CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
CC TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF
CC MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X80226; CAA56513.1; -.
CC InterPro; IPR002516; GT.11.
CC Pfam; PF01531; Glyco.transf.11; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 13 29 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).
FT DOMAIN 30 373 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 373 AA; 42098 MW; 0A47A1786231525C CRC64;

Query Match 53.5%; Score 1086.5; DB 1; Length 373;
Best Local Similarity 57.6%; Pred. No. 1.9e-84;
Matches 209; Conservative 51; Mismatches 62; Indels 41; Gaps 6;

QY 16 LIFVFT-----STIIHLQORIVK-----LQPLSEK-----LPMITQMS----- 50
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 LAFLLVCAISAPSFLLHLHODLFENGLALSPLCLERQPPVAPVAIVCLPVTSPASNASC 70

QY 51 SGNTESEPMRROSEHQHNGELRGMTFINSI GRIGNOMGEVATIFALARNGRGLAFIPAS 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 71 AGRPAAPS-----GIWTHPDGREGNMGQYATLLAQLNGRAFRATFQPEHMTTLAP 116

QY 111 HNALAPFRTISLPVLSHSDAKKIPWQNYHNDWMEERYRHIPGHVRFVTCYPCSWTFYH 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 117 HAALAPVFRITLPVLAPEVNRRTISWQLLLHDWMSSEYSLRLEDFPKLFTGFPSCWTFHH 176

QY 171 LRPEILKEFTLHDHVEEAQAFRLGRV--NGSQPSTFGVGVHVRGDIYHVMNVKGVV 228
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 177 VREQIRREFTLHDHVEEAQAFRLGRV--NGSQPSTFGVGVHVRGDIYHVMNVKGVV 236

QY 229 ADRGYLEKALDMFRARYSSPVFVTSNGMAWCENINASEGDIYHVRGDIYHVMNVKGVV 288
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 237 GDRAYLQQAQMDWFRAREAPFVYVTSNGMAWCENINASEGDIYHVRGDIYHVMNVKGVV 296

QY 289 LTOCNHTIMTIGTFGWAAYLAGDITTYLANYTLPDPSFLKFKPAAPFLPEWVGIPADL 348
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 297 LTOCNHTVMTIGTFGWAAYLAGDITTYLANYTLPDPSFLKFKPAAPFLPEWVGINADL 356

QY 349 SPL 351
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 357 SPV 359

RESULT 9
FUT1_HUMAN
ID FUT1_HUMAN STANDARD; PRT; 365 AA.
AC P19526;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)
DE (Fucosyltransferase 1) (Blood group H alpha 2-fucosyltransferase).
GN FUT1.
OS Homo sapiens (Human).
OX Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```



RA Meijerink E., Fries R., Voegeli P., Masabanda J., Wigger G.,  
 RA Stricker C., Neuenschwander S., Bertschinger H.U., Stranzinger G.,  
 RT "Two alpha(1,2) fucosyltransferase genes on porcine chromosome 6q11  
 RT are closely linked to the blood group inhibitor (S) and Escherichia  
 RT coli F18 receptor (ECF18R) loci.",  
 RL Mamm. Genome 8:736-741(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Small intestine;  
 RA Meijerink E., Neuenschwander S., Fries R., Dinter A.,  
 RA Bertschinger H.U., Stranzinger G., Voegeli P.,  
 RT "Expression and activity of porcine alpha(1,2)fucosyltransferases  
 RT determine erythrocyte antigen precursor O status and susceptibility to  
 RT Escherichia coli F18 colonization in the small intestine.",  
 RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +  
 CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.  
 CC -!- PATHWAY: Glycosylation.  
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND  
 CC FORM IN TRANS CISTERNAE OF GOLGI.  
 CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE  
 CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A  
 CC TISSUE-SPECIFIC MANNER.  
 CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; L50534; AAB02884.1; -  
 CC EMBL; U70883; AAB81884.1; -  
 CC EMBL; AF136896; AAF59833.1; -  
 CC InterPro; IPR002516; GT\_11.  
 CC Pfam; PF01531; Glyco\_transf\_11; 1.  
 CC Signal-anchor; Glycosyltransferase; Glycoprotein; Transmembrane;  
 CC Signal-anchor; Golgi stack; Polymorphism.  
 CC DOMAIN 1 8  
 CC TRANSMEM 9 28  
 CC DOMAIN 29 365  
 CC LUMENAL, ANCHOR (TYPE-II MEMBRANE PROTEIN).  
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 301 301  
 CC CARBOHYD 327 327  
 CC CARBOHYD 327 327  
 CC VARIANT 103 103  
 CC VARIANT 286 286  
 CC CONFLICT 77 77  
 CC SEQUENCE 365 AA; 41106 MW; DAFCE7E89A29D75 CRC64;  
 CC  
 CC Query Match 51.7%; Score 1050.5; DB 1; Length 365;  
 CC Best Local Similarity 62.9%; Pred. No. 2.1e-81;  
 CC Matches 195; Conservative 38; Mismatches 74; Indels 3; Gaps 2;  
 CC  
 CC QY 44 PMTQSSNTSPSPMRDSEHGNGELRGMTTINSIGRLGNOMGEYATLFLARNNGRL 103  
 CC DB 51 PVAIFCLAGTPVHPNASDSCPKH-PASLSGTWTYPDGRFGNQMGQYATLLALAQLNGRQ 109  
 CC QY 104 AFTPASMHNALAPIRISLPVLHSDAKKIPQNTYHLNDNMEERYEHIPGHVFTGYPC 163  
 CC DB 110 AFTQAMHVALPVPFRITLPLVLAPEVDHAPWRELELHDWMSDEYAHKPEKWKLTGFP 169  
 CC QY 164 SWTFVHLPEIKFTLHDHVREREAQFLGRV--NGSQSTFGVHVRRGDYVHVMP 221  
 CC DB 170 SWTFPHLREQIRSETLHDHLRQEAQVLSQFRLEPTGDRSTFGVHVRRGDYLRVMP 229  
 CC QY 222 NVWKGVDARGYLEKALDMFRARYSPFVVTNSGMWCRENINASRGDVFAGNGIEGS 281  
 CC DB 230 KRWKGVGDYLAQQAMDWFRARYAPFVVTNSGMWCRKNIDTSRGDSVIFAGDREAA 289  
 CC QY 282 PAKDFAALLTQCHNTIMTIGTFCINWAAAYLAGGDTIYLANVTLPSDFLKFVKPEAFPLPW 341  
 CC DB 282 PAKDFAALLTQCHNTIMTIGTFCINWAAAYLAGGDTIYLANVTLPSDFLKFVKPEAFPLPW 341

DB 290 PARDFALLVQCNTIMTIGTFCINWAAAYLAGGDTIYLANVTLPSDFLKFVKPEAFPLPW 349  
 QY 342 VGIADLSPL 351  
 DB 350 VGINADLSPL 359  
 RESULT 11  
 ID FUT2\_RAT  
 AC Q10984; STANDARD; PRT; 159 AA.  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (Secretor blood  
 DE group alpha-2-fucosyltransferase) (GDP-L-fucose-beta-D-  
 DE galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)FT 2)  
 DE (fucosyltransferase 2) (Fragment).  
 DE FUT2 OR SEC1 OR FTB.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BDIX;  
 RX MEDLINE=94280382; PubMed=8010942;  
 RA Piau J.-P., Labarriere N., Dabouis G., Denis M.G.;  
 RT "Evidence for two distinct alpha(1,2)-fucosyltransferase genes  
 RT differentially expressed throughout the rat colon.",  
 RL Biochem. J. 300:623-626(1994).  
 CC -!- FUNCTION: CREATES A MEMBRANE-ASSOCIATED PRECURSOR OLIGOSACCHARIDE  
 CC FUC-ALPHA(1,2)GALBETA-) CALLED THE H ANTIGEN WHICH IS AN  
 CC ESSENTIAL SUBSTRATE FOR THE FINAL STEP IN THE MEMBRANE-ASSOCIATED  
 CC A AND B ANTIGEN SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE  
 CC THE SAME ACCEPTOR SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.  
 CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +  
 CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.  
 CC -!- PATHWAY: Glycosylation.  
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND  
 CC FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).  
 CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE  
 CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A  
 CC TISSUE-SPECIFIC MANNER.  
 CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; L26010; AAB41515.1; -  
 CC InterPro; IPR002516; GT\_11.  
 CC Pfam; PF01531; Glyco\_transf\_11; 1.  
 CC Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane;  
 CC Golgi stack.  
 CC NON\_TER 1 1  
 CC CARBOHYD 98 98  
 CC SEQUENCE 159 AA; 17250 MW; 14DECEB7C2E6384A CRC64;  
 CC  
 CC Query Match 41.6%; Score 846; DB 1; Length 159;  
 CC Best Local Similarity 99.4%; Pred. No. 1.5e-64;  
 CC Matches 158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC QY 222 NVWKGVDARGYLEKALDMFRARYSPFVVTNSGMWCRENINASRGDVFAGNGIEGS 281  
 CC DB 1 NVWKGVDARGYLEKALDMFRARYSPFVVTNSGMWCRENINASRGDVFAGNGIEGS 60  
 CC QY 282 PAKDFAALLTQCHNTIMTIGTFCINWAAAYLAGGDTIYLANVTLPSDFLKFVKPEAFPLPW 341  
 CC DB 282 PAKDFAALLTQCHNTIMTIGTFCINWAAAYLAGGDTIYLANVTLPSDFLKFVKPEAFPLPW 341

Db 173 ---DKYKEKYSQSVKPAIDFSPLEPNSDSNFISGICHIHRTDFVDGQHSSNV----- 223

QY 230 DRGYLEKALDMFRAR-----YSPFVVTNSGMAWRENINASRGD 270

Db 224 --SEIKPALEIFEREQKDYNKMLIVIMGDDDFEAKMF-----EGTVRAKKEAKLEET 277

QY 271 VVFAGNIEGSPAKDFALL-TCGNHTIMTI--GTGTWAIYLAGGDTIYLANV-TLPDSP 326

Db 278 KYEVS---ENIPQDDLAYSHYSCDVTILTPASTFTGWMGLGYSKGQAVIYQDIRSINDYN 334

QY 327 FLK-VFKPEAAFLPEWVGIPAD 347

Db 335 YKKGVLDPDDFFVPSWTSIMLD 356

RESULT 13

ID	VGLD_HSVEK	STANDARD;	PRT;	442 AA.
AC	P22484;			
DT	01-AUG-1991 (Rel. 19, Created)			
DT	01-AUG-1991 (Rel. 19, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Glycoprotein D precursor (Glycoprotein 17/18).			
GD	OR GP17/18 OR 72.			
OS	Equine herpesvirus type 1 (strain Kentucky A) (EHV-1).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Alphaherpesvirinae; Varicellovirus.			
OX	NCBI_TaxID=10329;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=91082407; PubMed=1845821;			
RA	Flowers C.C., Eastman E.M., O'Callaghan D.J.;			
RT	"Sequence analysis of a glycoprotein D gene homolog within the unique			
RT	short segment of the EHV-1 genome.";			
RL	Virology 180:175-184(1991).			
RP	[2]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=92263758; PubMed=1316673;			
RA	Colle C.F., III, Flowers C.C., O'Callaghan D.J.;			
RT	"Open reading frames encoding a protein kinase, homolog of			
RT	glycoprotein gx of pseudorabies virus, and a novel glycoprotein map			
RT	within the unique short segment of equine herpesvirus type 1.";			
RL	Virology 188:545-557(1992).			
CC	!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN D FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collabor			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstat			
CC	the European Bioinformatics Institute. There are no restrictions on			
CC	use by non-profit institutions as long as its content is in no			
CC	modified and this statement is not removed. Usage by and for comm			
CC	entities requires a license agreement (See http://www.isb-sib.ch/anno			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL; M62923; AAA46081.1; -			
DR	EMBL; M66931; -; NOT_ANNOTATED_CDS.			
DR	EMBL; M67497; AAA46073.1; ALT_INIT.			
DR	PIR; A36518; VGBSEA.			
DR	InterPro; IPR002896; Herpes_glycop.D.			
DR	Pfam; PF01537; Herpes_glycop.D; 1.			
KW	Glycoprotein; Signal; Transmembrane.			
FT	SIGNAL 1 19 POTENTIAL.			
FT	CHAIN 20 442 GLYCOPROTEIN D.			
FT	DOMAIN 20 405 EXTRACELLULAR (POTENTIAL).			
FT	TRANSMEM 406 422 POTENTIAL.			
FT	DOMAIN 423 442 CYTOPLASMIC (POTENTIAL).			
FT	CARBOHYD 103 103 N-LINKED (GLCNAC. .) (POTENTIAL).			
FT	CARBOHYD 111 111 N-LINKED (GLCNAC. .) (POTENTIAL).			
FT	CARBOHYD 347 347 N-LINKED (GLCNAC. .) (POTENTIAL).			
FT	CARBOHYD 396 396 N-LINKED (GLCNAC. .) (POTENTIAL).			
SEQ	SEQUENCE 442 AA; 49908 MW; 323CDCA9C9762F05 ORG64;			

Query Match 4.5%; Score 91; DB 1; Length 442;  
 Best Local Similarity 22.7%; Pred. No. 3.2;



Query Match	4.4%; Score 89; DB 1; Length 353;
Best Local Similarity	22.1%; Pred. No. 3.6;
Matches	64; Conservative 38; Mismatches 110; Indels 78; Gaps 12;

  

QY	45	MTTCSSGNTESPMDRDSQ-----HGNGELRGMTTINSIGRLGNQMG---EYALIFA 95
DB	80	LTTELTEGNEAGLIMTABFEVTLVAGN-----NYGYDGSQGNQINSQGTPL 130
QY	96	LARNGLGLAF--IPASMHNALAPIFRISLPVLHSDTAKK-----132
DB	131	IKRVHARIATKTEVTSQSVANKYNFAPENIVALYAKKSNLPGASLANSDDAYLTGSL 190
QY	133	-----IPWONYHLNDWMEERYRHI-----PGHFYRFTGYPCSWTFYHH--LRPEIL- 176
DB	191	TTENGAVSPANYTHVDWLGRDYTEIGAATVTPKGF-----YVLESTYAGNAGLRPTILC 245
QY	177	--KEFTLHDHV---REEAQAFLGLVNSQSSTPF--VGVHVRGDIYVHMNVKGVV 228
DB	246	VKGKLTREDGTALSSSEMTAFNAGWIVANNDPTTYPIVLNPFESNNYTYTGEAVEKGKI 305
QY	229	ADRGYLEKALDMPFARYSSPVVYTSNGMAWCINASRGDVVFAGNGI 278
DB	306	VRNHKFDINTITGPTGNPENPIT-----ESANLNVCVVAANKGV 347

  

RESULT 15				
CDH_PHACH				
ID	CDH_PHACH	STANDARD;	PRT;	773 AA.
AC	Q01738;	Q00047;		
DT	30-MAY-2000	(Rel. 39, Created)		
DT	30-MAY-2000	(Rel. 39, Last sequence update)		
DT	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	Cellulose dehydrogenase precursor (EC 1.1.5.1) (CDH) (Cellobiose-			
DE	quinone oxidoreductase).			
GN	CDH-1 AND CDH-2.			
CS	Phanerochaete chrysosporium.			
OC	Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;			
OC	Aphyllophorales; Corticiales; Phanerochaete.			
OX	NCBI_TaxID=5306;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=OGC101;			
RX	MEDLINE=97077226; PubMed=8919793;			
RA	Li B., Nagalla S.R., Renganathan V.;			
RT	"Cloning of a cDNA encoding Cellobiose dehydrogenase, a			
RT	hemoflavoenzyme from Phanerochaete chrysosporium.";			
RL	Appl. Environ. Microbiol. 62:1329-1335(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=OGC101;			
RX	MEDLINE=97176414; PubMed=9023960;			
RA	Li B., Nagalla S.R., Renganathan V.;			
RT	"Cellulose dehydrogenase from Phanerochaete chrysosporium is encoded			
RT	by two allelic variants.";			
RL	Appl. Environ. Microbiol. 63:796-799(1997).			
RN	[3]			
RP	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 19-208.			
RC	MEDLINE=20139694; PubMed=10673428;			
RA	Hallberg B.M., Bergfors T., Boeckbro K., Pettersson G., Henriksson G.,			
RA	Divne C.;			
RT	"A new scaffold for binding haem in the cytochrome domain of the			
RT	extracellular flavocytochrome cellobiose dehydrogenase.";			
RL	Structure 8:79-88(2000).			
CC	-1-	FUNCTION: DEGRADERS BOTH LIGNIN AND CELLULOSE. OXIDIZES CELLOBIOSE		
CC	CC	TO CELLOBIONOLACTONE.		
CC	-1-	CATALYTIC ACTIVITY: Cellobiose + a quinone = cellobiono-1,5-		
CC	CC	lactone + a phenol.		
CC	-1-	COFACTOR: ONE FAD AND ONE HEME B.		
CC	-1-	SUBCELLULAR LOCATION: Secreted		
CC	-1-	SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE GMC		
CC	CC	OXIDOREDUCTASES FAMILY.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			



GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:52:27 ; Search time 45.4049 Seconds  
(without alignments)  
1050.630 Million cell updates/sec

Title: US-10-040-863-8\_COPY\_23\_380

Perfect score: 1921

Sequence: 1 STIIHQRIIVKLPLSEKE.....RSHFLKAKGVTCYVAGRAF 358

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1921	100.0	380	21	Rat hepatoma H35 C
2	1921	100.0	380	23	Rat hepatoma H35 C
3	1896	98.7	353	21	Rat hepatoma H35 C
4	1896	98.7	353	23	Rat hepatoma H35 C
5	1451	75.5	340	19	Porcine secretor t
6	1451	75.5	340	19	Pig secretor. Sus
7	1407.5	73.3	343	19	Human Sec2 protein
8	1397	72.7	344	21	Human Sec2 catalyt
9	1051.5	54.7	365	12	GDP-Fuc:beta-D-gal
10	1051.5	54.7	365	16	GDP-L-fucose-beta-

11	1051.5	54.7	365	16	2-Alpha-fucosyltra
12	1051.5	54.7	365	16	2-Alpha-fucosyltra
13	1051.5	54.7	365	17	Human H-transferas
14	1051.5	54.7	365	18	Human alpha 1,2-fu
15	1051.5	54.7	365	18	Human alpha(1,2)-f
16	1051.5	54.7	365	21	Human H-transferas
17	1046.5	54.5	365	19	Pig H transferase.
18	1046.5	54.5	365	19	Pig alpha-1-2 fucco
19	1046.5	54.5	365	23	Swine alpha(1,2)
20	1044.5	54.4	365	20	Swine alpha(1,2) f
21	1044.5	54.4	365	20	Swine alpha(1,2) f
22	1043.5	54.3	365	15	A glycosyltransfer
23	990	51.5	357	20	X. laevis alpha-1,
24	288.5	15.0	110	21	Human secreted pro
25	141	7.3	300	21	Human secreted pro
26	141	7.3	300	21	AA192713
27	139.5	7.3	287	23	AA192713
28	103	5.4	661	22	AA192713
29	101	5.3	690	23	AA192713
30	97	5.0	317	13	AA192713
31	96.5	5.0	355	21	AA192713
32	95	4.9	1515	22	AA192713
33	93	4.8	1073	22	AA192713
34	92	4.8	984	22	AA192713
35	89.5	4.7	586	21	AA192713
36	89.5	4.7	604	22	AA192713
37	89	4.6	353	16	AA192713
38	89	4.6	394	22	AA192713
39	89	4.6	1031	22	AA192713
40	89	4.6	2057	22	AA192713
41	89	4.6	2058	23	AA192713
42	88.5	4.6	1416	22	AA192713
43	88	4.6	217	23	AA192713
44	88	4.6	629	22	AA192713
45	87	4.5	217	16	AA192713

#### ALIGNMENTS

#### RESULT 1

AA192713

ID AAB36104 standard; Protein; 380 AA.

XX

AC AAB36104;

XX

DT 19-FEB-2001 (first entry)

XX

DE Rat hepatoma H35 cell alpha1-2fucosyltransferase.

XX

KW Rat: alpha1-2fucosyltransferase; cytostatic; neuroprotective;

KW notropic; gene therapy; Fucalalpha1-2galbeta1-3galNAc; immunotherapy;

KW immunosuppression; cancer; neurological disease;

KW small cell lung carcinoma.

XX

OS Rattus norvegicus.

XX

PN WO2000064464-A1.

XX

PD 02-NOV-2000.

XX

PF 23-APR-1999; 99WO-US07384.

XX

PR 23-APR-1999; 99WO-US07384.

XX

PA (PACI-) PACIFIC NORTHWEST CANCER FOUND.

XX

PI Holmes EH, Sherwood AL;

XX

DR WPI: 2000-687262/67.

DR N-PSDB; AAC67965.

XX

PT New rat ganglioside GM1-specific alpha1-2fucosyltransferase, useful for



```

Db 323 PDSFPLKVKPEAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 380
RESULT 3
ID AAB36105 standard; Protein; 353 AA.
XX AAB36105;
XX 19-FEB-2001 (first entry)
XX Rat hepatoma H35 cell alpha1-2fucosyltransferase catalytic domain.
XX Rat; alpha1-2fucosyltransferase; cytostatic; neuroprotective;
KW neotropic; gene therapy; Fucalpal-2Galbeta1-3GalNAc; immunotherapy;
KW immunosuppression; cancer; neurological disease;
KW small cell lung carcinoma.
XX Rattus norvegicus.
XX WC2000064464-A1.
XX 02-NOV-2000.
XX 23-APR-1999; 99WO-US07384.
XX 23-APR-1999; 99WO-US07384.
XX (PACT-) PACIFIC NORTHWEST CANCER FOUND.
XX Holmes EH, Sherwood AL;
XX WPI; 2000-687262/67.
XX N-PSDB; AAC67966.
XX New rat ganglioside GM1-specific alpha1-2fucosyltransferase, useful for
PT preparation of fucosyl GM1 which is useful as a nutritional composition
PT or immunotherapeutic for cancer and neurological diseases -
XX Claim 2; Fig 3A; 91pp; English.
XX The present sequence is given in a specification relating to a rat
CC ganglioside GM1-specific alpha1-2fucosyltransferase protein. The protein
CC or its cellular fraction is useful for synthesis of a molecule comprising
CC Fucalpal-2Galbeta1-3GalNAc, a glycolipid, glycoprotein, glycolipoprotein
CC or a free oligosaccharide comprising Fucalpal-2Galbeta1-3GalNAc.
CC The method involves contacting alpha1-2fucosyltransferase with GDP-fucose
CC and a molecule or glycolipid, glycoprotein, glycolipoprotein or
CC oligosaccharide having a terminal Galbeta1-3GalNAc group. It is also
CC useful for synthesis of fucosyl-GM1 by contacting the protein with
CC GDP-fucose and ganglioside GM1. The obtained glycoproteins,
CC glycolipoproteins, glycolipids and oligosaccharides are useful as
CC nutritional compositions and fucosyl-GM1 is useful for inducing an
CC immunotherapeutic or immunosuppressive action against cancer,
CC neurological disease or small cell lung carcinoma.
XX Sequence 353 AA;
Query Match 98.7%; Score 1896; DB 21; Length 353;
Best Local Similarity 100.0%; Pred. No. 1.3e-198;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 LQORTVKLOPSEKELPMTTQSSNGTSPERRDSEQHNGELRGMTINSIGRLNQM 65
DB 1 LQORTVKLOPSEKELPMTTQSSNGTSPERRDSEQHNGELRGMTINSIGRLNQM 60
QY 66 GEYATLALARNGLATIPASMHNALAPIFRISLPVLHSDTAKKIPQWYHLNDWMEER 125
DB 61 GEYATLALARNGLATIPASMHNALAPIFRISLPVLHSDTAKKIPQWYHLNDWMEER 120
QY 126 YRHIPGHVVRTGYPCSWTFYHHLRPEILKEFTLHDHVRBAQAFLRLVNGSQPSTFV 185
DB 121 YRHIPGHVVRTGYPCSWTFYHHLRPEILKEFTLHDHVRBAQAFLRLVNGSQPSTFV 180
QY 186 GVHVRGDIYHVMPNVKGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCRENINAS 245
DB 181 GVHVRGDIYHVMPNVKGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCRENINAS 240
QY 246 RGDVVFAGNGIEGSPAKDFALLTQCNTIMTIGTGWAAAYLAGGDTIYLYANTLPDPPF 305
DB 241 RGDVVFAGNGIEGSPAKDFALLTQCNTIMTIGTGWAAAYLAGGDTIYLYANTLPDPPF 300
QY 306 LKVFKEPAALPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 358
DB 301 LKVFKEPAALPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 353
RESULT 4
AAEL16623
ID AAEL16623 standard; Protein; 353 AA.
XX AAEL16623;
XX 09-APR-2002 (first entry)
XX Rat hepatoma H35 cell alpha1-2Fuct catalytic domain.
XX Rat; alpha1-2fucosyltransferase; alpha1-2Fuct; antisense therapy;
KW galactose beta1-3N-acetylgalactosamine; Galbeta1-3GalNAc; glycolipid;
KW glycoprotein; glycolipoprotein; oligosaccharide; fucosyl-GM1; cancer;
KW gene therapy; oncogenic transformation; cytostatic; ganglioside; GM1;
KW cell transformation; catalytic domain.
XX Rattus norvegicus.
XX Key Location/Qualifiers
FH Modified-site 172
FT /note= "N-glycosylated"
FT Region 194..353
FT /note= "Region which overlaps rat FTB"
FT Modified-site 238
FT /note= "N-glycosylated"
FT Modified-site 266
FT /note= "N-glycosylated"
FT Modified-site 292
FT /note= "N-glycosylated"
XX US6329170-B1.
XX 11-DEC-2001.
XX 23-APR-1999; 99US-0298886.
XX 23-APR-1999; 99US-0298886.
XX (NMHO-) NORTHWEST HOSPITAL.
XX Holmes EH, Sherwood AL;
XX WPI; 2002-121132/16.
XX N-PSDB; AAD27206.
XX Rat hepatoma H35 cell alpha1-2fucosyltransferase, useful for producing
PT GM1-specific alpha1-2fucosyltransferase enzyme by recombinant
PT techniques and for detecting oncogenic transformation of test tissues -
XX Claim 4; Fig 3; 41pp; English.
XX The invention relates to rat GM1-specific alpha1-2fucosyltransferase
CC (alpha1-2Fuct) enzyme and its corresponding nucleic acid. This nucleic
CC acid is specific for a carbohydrate moiety found in ganglioside GM1,
CC a terminal galactose beta1-3N-acetylgalactosamine (Galbeta1-3GalNAc)
CC saccharide. Alpha1-2Fuct DNA is useful for producing rat alpha1-2Fuct
CC protein by recombinant techniques. Alpha1-2Fuct DNA is useful for the
CC preparative synthesis of fucosyl containing glycolipids, glycoproteins,
CC glycolipoproteins and oligosaccharide, and for preparing fucosyl-GM1.

```

CC Alpha1-2Fuct DNA is useful for detecting oncogenic transformation which  
 CC involves assaying for changes in expression of alpha1-2 Fuct. Since  
 CC alpha1-2Fuct is activated in cell transformation, antisense sequences  
 CC derived from alpha1-2Fuct DNA are useful for inhibiting, suppressing  
 CC or treating cancer. Alpha1-2Fuct DNA is useful in gene therapy and  
 CC antisense therapy. The present sequence is rat hepatoma H35 cell  
 CC alpha1-2Fuct catalytic domain.  
 XX  
 SQ Sequence 353 AA;

Query Match 98.7%; Score 1896; DB 23; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-198;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 LQORIVKLQPLSEKELPMTTOMSSGNTSEPMRRDSEQHNGELRGMTINSIGRLGNQM 65  
 DB 1 LQORIVKLQPLSEKELPMTTOMSSGNTSEPMRRDSEQHNGELRGMTINSIGRLGNQM 60  
 QY 56 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER 125  
 DB 61 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER 120  
 QY 126 YRHIPGSHVFRFTGYPSCSWTFYHHLRPEILKFTLLHDHVRERAAQAFILGLRVNGSOPSTFV 185  
 DB 121 YRHIPGSHVFRFTGYPSCSWTFYHHLRPEILKFTLLHDHVRERAAQAFILGLRVNGSOPSTFV 180  
 QY 186 GVHVRGSDYVHVMFNWVGVDAGYLYLEKALDMFRARYSSPVVVTSGMAWCRENINAS 245  
 DB 181 GVHVRGSDYVHVMFNWVGVDAGYLYLEKALDMFRARYSSPVVVTSGMAWCRENINAS 240  
 QY 246 RGVVVFAGNIEGSAKDFALLTQCNNHIMIGTPIGIIWAAYLAGGDTIYLANIYLPDSPF 305  
 DB 241 RGVVVFAGNIEGSAKDFALLTQCNNHIMIGTPIGIIWAAYLAGGDTIYLANIYLPDSPF 300  
 QY 306 LKVPKPEAAFLPEWVGIPADLSPLKALTPACPRSHFHLKAKGVTCYVAGRAF 358  
 DB 301 LKVPKPEAAFLPEWVGIPADLSPLKALTPACPRSHFHLKAKGVTCYVAGRAF 353

RESULT 5  
 AAW37855  
 ID AAW37855 standard; Protein; 340 AA.

XX AC AAW37855;

XX DT 28-AUG-1998 (first entry)

XX DE Porcine secretor transferase (FUT2).

XX KW Secretor; glycosyltransferase; FUT2; pig; epitope; antigen;  
 KW transgenic animal; xenotransplantation; organ transplant.

XX OS Sus scrofa.

XX FH Key Location/Qualifiers

FT Domain 1..4

FT /note= "N-terminal cytoplasmic tail"

FT Domain 5..26

FT /note= "transmembrane domain"

FT Domain 27..340

FT /note= "C-terminal domain"

FT Modified-site 185..187

FT /note= "Asn is N-glycosylated"

FT Modified-site 251..253

FT /note= "Asn is N-glycosylated"

FT Modified-site 279..281

FT /note= "Asn is N-glycosylated"

FT Modified-site 305..307

FT /note= "Asn is N-glycosylated"

XX WO9807837-A1.

XX PN 26-FEB-1998.

XX AC AAW53101;

XX 22-AUG-1997; 97WO-AU00540.  
 XX PR 23-AUG-1996; 96AU-0001823.  
 XX (AUST-) AUSTIN RES INST.  
 XX McKenzie IFC, Sandrin MS;  
 DR WPI; 1998-169148/15.  
 DR N-PSDB; AAV29003.  
 PT Nucleic acid encoding glycosyltransferase able to compete with  
 PT second such enzyme - particularly used to reduce expression of  
 PT unwanted carbohydrate epitope(s) on tissues intended for  
 PT transplantation  
 XX  
 PS Claim 6; Fig 1A-B; 40pp; English.

XX This polypeptide comprises porcine secretor glycosyltransferase  
 CC (SE or FUT2), a type II integral membrane protein has high affinity  
 CC for type I and type III substrates. Its amino acid sequence was  
 CC deduced from the nucleotide sequence of a genomic DNA clone (see  
 CC AAV29003) isolated from a pig liver library on the basis of homology  
 CC to human FUT2 cDNA. Pig FUT2 shows 83.2% identity with human FUT2.  
 CC Expression of a glycosyltransferase, such as porcine Se, in a  
 CC tissue results in reduced expression of unwanted carbohydrate  
 CC epitopes on the tissue, especially porcine heart, liver, kidney or  
 CC pancreas, rendering it more suitable for transplantation, i.e. less  
 CC immunogenic and of increased immunological acceptability. A  
 CC claimed method of producing a cell from a donor species that is  
 CC immunologically acceptable to a recipient species involves reducing  
 CC levels of carbohydrate on the donor cell that causes it to be  
 CC recognised as non-self by the recipient by expressing a nucleic  
 CC acid for a glycosyltransferase such as porcine Se in the cell.

XX SQ Sequence 340 AA;

Query Match 75.5%; Score 1451; DB 19; Length 340;  
 Best Local Similarity 81.9%; Pred. No. 6.8e-150;  
 Matches 271; Conservative 26; Mismatches 20; Indels 14; Gaps 3;

QY 1 STIIHLQORIVKVLQPLSEKELPMTTOMSSGNTSEPMRRDSEQHNGELRGMTINSIGR 60  
 DB 23 STIHLQORIVKVLQPLSEKELPMTTOMSSGNTSEPMRRDSEQHNGELRGMTINSIGR 68  
 QY 61 LGNQMGVATLFLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLND 120  
 DB 59 LGNQMGVATLFLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLND 128  
 QY 121 WMEERYRHIPGSHVFRFTGYPSCSWTFYHHLRPEILKFTLLHDHVRERAAQAFILGLRVNSQ 180  
 DB 129 WMEERYRHIPGSHVFRFTGYPSCSWTFYHHLRPEILKFTLLHDHVRERAAQAFILGLRVNSQ 188  
 QY 181 PSTFVGVHVRGSDYVHVMFNWVGVDAGYLYLEKALDMFRARYSSPVVVTSGMAWCRE 240  
 DB 189 PSTFVGVHVRGSDYVHVMFNWVGVDAGYLYLEKALDMFRARYSSPVVVTSGMAWCRE 248  
 QY 241 NINASRGDVPVAGNIEGSAKDFALLTQCNNHIMIGTPIGIIWAAYLAGGDTIYLANIYTL 300  
 DB 249 NINASRGDVPVAGNIEGSAKDFALLTQCNNHIMIGTPIGIIWAAYLAGGDTIYLANIYTL 308  
 QY 301 PDSPLKLVKFKPEAAFLPEWVGIPADLSPLK 331  
 DB 309 PDSPLKLVKFKPEAAFLPEWVGIPADLSPLK 339

RESULT 6  
 AAW53101.  
 ID AAW53101 standard; Protein; 340 AA.

```
DT 08-JUL-1998 (first entry)
XX
DE Pig secretor.
XX
KW Pig; secretor; chimeric; glycosyltransferase; gene therapy;
KW transplantation.
XX
OS Sus scrofa.
XX
PN WC9805768-Al.
XX
PD -12-FEB-1998.
XX
XX 01-AUG-1997; 97WO-AU00492.
XX
XX 21-AUG-1996; 96US-0024279.
XX
XX 02-AUG-1996; 96AU-0001402.
XX
XX (AUST-) AUSTIN RES INST.
XX
XX McKenzie IFC, Sandrin MS;
XX
XX WPI; 1998-159170/14.
XX
XX N-PSDB; AAV21639.
XX
XX Nucleic acids encoding chimeric glycosyltransferases - used for
XX PT altering carbohydrate levels on the surface of cells, useful in gene
XX PT therapy and transplantation
XX
XX Example 2; Fig 6; 51pp; English.
XX
XX The present sequence represents pig secretor used in an example of the
XX CC present invention. The present invention describes nucleic acids (NA)
XX CC encoding a chimeric glycosyltransferase. The NA comprises a catalytic
XX CC domain of a first glycosyltransferase (GT) and a localisation signal of
XX CC a second GT, whereby when the NA is expressed in a cell and where the
XX CC chimeric enzyme is located in an area of the cell where it is able to
XX CC compete for substrate with a second GT, resulting in reduced levels of
XX CC a product from the second GT. The NA can be used to produce cells and
XX CC organs with desired glycosylation patterns. Products and methods of the
XX CC present invention can be used to reduce the levels of undesirable
XX CC epitopes in cells, tissues or organs which may be used in
XX CC transplantation or gene therapy.
XX
XX Sequence 340 AA;
XX
XX Query Match 75.5%; Score 1451; DB 19; Length 340;
XX Best Local Similarity 81.9%; Pred. No. 6.8e-150;
XX Matches 271; Conservative 26; Mismatches 20; Indels 14; Gaps 3;
XX
XX QY 1 STTIIHQRIKVLQPLSEKELPMTTOMSSGNTSPSPMRDSECHGNGELRGMTTINSIGR 60
XX Db 23 STIFHLQQRWVKIQP--TWELQWVTQV---TESP-----SSPQLKGMWTINAIGR 68
XX
XX QY 61 LGNOMGEYATLALARMNGRLAFIPASMNALAPIRISLPVLHSDTAKKIPQWNYHLND 120
XX Db 69 LGNOMGEYATLALARMNGRPAFIPPEMHSTLAPIRITIPVLHASTARRIPQWNYHLND 128
XX
XX QY 121 WMEERYRHIPGHVFRVTGPGCSWTFYHHLRPELTKETLHDHVREEAQAFLRGLRYNGSQ 180
XX Db 129 WMEERYRHIPGHVFRVTGPGCSWTFYHHLRPELTKETLHDHVREEAQAFLRGLRYNGSR 188
XX
XX QY 181 PSTFVGHVHVRGDIYHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTNSNGMAWCRE 240
XX Db 189 PSTFVGHVHVRGDIYHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTNSNGMAWCRE 248
XX
XX QY 241 NINASRGDVVFAGNTEGSPAKDFALLTCQNTHTWITGTFGIWAAYLAGSDTIYLANLYL 300
XX Db 249 NINASRGDVVFAGNTEGSPAKDFALLTCQNTHTWITGTFGIWAAYLAGSDTIYLANLYL 308
XX
XX QY 301 PDSFPLKFKPEAAFLPENWGPADLSPLLK 331
XX Db 309 PDSFPLKFKPEAAFLPEWIGTEADLSPLLK 339
```

```
RESULT 7
AAW69332
ID AAW69332 standard; Protein; 343 AA.
XX
AC AAW69332;
XX
DT 20-NOV-1998 (first entry)
XX
DE Human Sec2 protein sequence.
XX
XX Sec2; alpha(1,2) fucosyltransferase; H blood group; secretor genotyping;
XX KW GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase; human;
XX KW FUT2; nonsecretor genotyping.
XX
XX OS Homo sapiens.
XX
XX US5807732-A.
XX
XX PN 15-SEP-1998.
XX
XX PD 28-FEB-1995; 95US-0395800.
XX
XX PF 28-FEB-1995; 95US-0395800.
XX
XX PR 28-FEB-1995; 95US-0395800.
XX
XX (GIOR/) GIORGI D.
XX
XX (KELL/) KELLY R J.
XX
XX (LENN/) LENNON G.
XX
XX (LOWE/) LOWE J B.
XX
XX (ROUQ/) ROUQUIER S.
XX
XX Giorgi D, Kelly RJ, Lennon G, Lowe JB, Rouquier S;
XX WPI; 1998-320127/44.
XX
XX N-PSDB; AAV58323.
XX
XX DNA encoding fucosyltransferase enzyme - useful for producing
XX PT recombinant enzyme and genotyping person as secretor or nonsecretor
XX
XX PS Disclosure; Column 45-50; 55pp; English.
XX
XX This sequence is the human Sec2 protein of the invention. The DNA
XX CC encodes a alpha(1,2) fucosyltransferase and is the secretor
XX CC alpha(1,2)fucosyltransferase locus that cross hybridises with the
XX CC H blood group alpha(1,2)fucosyltransferase gene. The DNA is useful for
XX CC producing a recombinant human GDP-L-fucose:beta-D-galactoside
XX CC 2-alpha-L-fucosyltransferase (FUT2) which can be used for genotyping an
XX CC individual as a secretor or nonsecretor as it is known that nonsecretors
XX CC homozygous for a mutant allele of the FUT2 gene that has a stop codon in
XX CC the position corresponding to amino acid 143.
XX
XX Sequence 343 AA;
XX
XX Query Match 73.3%; Score 1407.5; DB 19; Length 343;
XX Best Local Similarity 79.2%; Pred. No. 4e-145;
XX Matches 262; Conservative 26; Mismatches 32; Indels 11; Gaps 2;
XX
XX QY 1 STTIIHQRIKVLQPLSEKELPMTTOMSSGNTSPSPMRDSECHGNGELRGMTTINSIGR 60
XX Db 23 STIFHVQQRILAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGMWTINAIGR 71
XX
XX QY 61 LGNOMGEYATLALARMNGRLAFIPASMNALAPIRISLPVLHSDTAKKIPQWNYHLND 120
XX Db 72 LGNOMGEYATLALARMNGRPAFIPACMHSTLAPIRITIPVLHASTARRIPQWNYHLND 131
XX
XX QY 121 WMEERYRHIPGHVFRVTGPGCSWTFYHHLRPELTKETLHDHVREEAQAFLRGLRYNGSQ 180
XX Db 132 WMEERYRHIPGHVFRVTGPGCSWTFYHHLRPELTKETLHDHVREEAQAFLRGLRYNGSR 191
XX
XX QY 181 PSTFVGHVHVRGDIYHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTNSNGMAWCRE 240
XX Db 192 PGTFTVGHVHVRGDIYHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTNSNGMAWCRE 251
```

QY 61 LGNQMEYATLFALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWONYHLND 120



Db 59 GTAMGNASSCPQH-PASLSGTWTVPNGRFGNQGYATLALALQNGRRAFILPAMH 117  
 QY 90 NALAPIFRISLPVLHSDIAKKIPWQNYHLNDWMEERYRHPGHEVFTGYPCSWTFYHHL 149  
 Db 118 AALAPVFRITLPLVLAPEVDSRTPWRQLQHDWSEYADLRDPFLKLSGFPSCSWTFHHL 177  
 QY 150 REELIKEFTLHDHVRBEAQAFLRGLRV--NGSQPSTFVGHVRRGDYVHVMPNWKGVYA 207  
 Db 178 REQIRREFTLHDHLEEAQSVLGQLRGTGDRPTFTVGHVRRGDYLVQVMPQKGVVG 237  
 QY 208 DRGYLEKALDMFRARYSSPVFVYVTSNGMAWCENINASRGDVVFAGNGIEGSPAKDFALL 267  
 Db 238 DSAYLRQAMDWFRARHEARVFFVTSNGMEWCENIDTSQGDVTFAGDGEATPWKDFALL 297  
 QY 268 TCNHTTMTIGTGFIAWYLAGDITIYLANFTLPDSEFLKFKPEAAFLPEWVGIPADLS 327  
 Db 298 TCNHTTMTIGTGFIAWYLAGDITIYLANFTLPDSEFLKFKPEAAFLPEWVGINADLS 357  
 QY 328 PLLKALTP 335  
 Db 358 PLWTLAKP 365

RESULT 10  
 AAR80154  
 ID AAR80154 standard; Protein; 365 AA.  
 XX AC AAR80154;  
 XX DT 25-JUN-1996 (first entry)  
 XX DE GDP-L-fucose-beta-D-galactosidase alpha-1,2-fucosyltransferase.  
 XX KW GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyl-transferase; Fuc-T;  
 KW synthesis; 2'-fucosyllactose; oligosaccharide; human milk;  
 KW non-human transgenic mammal; secondary gene product.  
 XX OS Homo sapiens.  
 XX PN WO9524494-Al.  
 XX PD 14-SEP-1995.  
 XX PF 24-JAN-1995; 95WO-US00926.  
 XX PR 09-MAR-1994; 94US-0209122.  
 XX PA (ABBO ) ABBOTT LAB.  
 XX PI Cummings RD, Kopchik JJ, Moreman KW, Mukerji P;  
 PI Pierce JM, Prieto PA, Smith DF;  
 XX DR WPI; 1995-336739/43.  
 XX DR N-PSDB; AAR98461.  
 XX PT Prodn. of humanised milk by non-human transgenic mammal - by  
 PT inserting heterologous gene encoding human catalytic entity which  
 PT produces oligo:saccharide(s) and glyco:conjugate(s) into mammal's  
 PT genome  
 XX PS Example 1; Page 62-64; 83pp; English.  
 XX CC GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyl-transferase (Fuc-T)  
 CC is encoded by cDNA (AAQ98461) isolated from a human epidermal carcinoma  
 CC cell line. The enzyme is responsible for the synthesis of  
 CC 2'-fucosyllactose, one of the oligosaccharides in human milk. Other  
 CC products of Fuc-T include glycoproteins contg. beta-linked terminal  
 CC galactose residues which can be fucosylated by Fuc-T. This DNA can  
 CC be used to produce non-human transgenic mammals able to produce  
 CC secondary gene products, e.g. oligosaccharides, in their milk. The  
 CC transgenic mammals milk biochemically resembles human milk. This  
 CC humanised milk can be used in the prepn. of an enteral nutritional  
 CC product useful in the nutritive maintenance of an animal.

XX SQ Sequence 365 AA;  
 Query Match 54.7%; Score 1051.5; DB 16; Length 365;  
 Best Local Similarity 63.3%; Pred. No. 4.3e-106;  
 Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;  
 QY 30 GNIESPEMRDRSDSQHNGELRGMTFINSIGRLGNQKGEYATLEALAPMNGRIAFIPASME 89  
 Db 59 GTAMGNASSCPQH-PASLSGTWTVPNGRFGNQGYATLALALQNGRRAFILPAMH 117  
 QY 90 NALAPIFRISLPVLHSDIAKKIPWQNYHLNDWMEERYRHPGHEVFTGYPCSWTFYHHL 149  
 Db 118 AALAPVFRITLPLVLAPEVDSRTPWRQLQHDWSEYADLRDPFLKLSGFPSCSWTFHHL 177  
 QY 150 REELIKEFTLHDHVRBEAQAFLRGLRV--NGSQPSTFVGHVRRGDYVHVMPNWKGVYA 207  
 Db 178 REQIRREFTLHDHLEEAQSVLGQLRGTGDRPTFTVGHVRRGDYLVQVMPQKGVVG 237  
 QY 208 DRGYLEKALDMFRARYSSPVFVYVTSNGMAWCENINASRGDVVFAGNGIEGSPAKDFALL 267  
 Db 238 DSAYLRQAMDWFRARHEARVFFVTSNGMEWCENIDTSQGDVTFAGDGEATPWKDFALL 297  
 QY 268 TCNHTTMTIGTGFIAWYLAGDITIYLANFTLPDSEFLKFKPEAAFLPEWVGIPADLS 327  
 Db 298 TCNHTTMTIGTGFIAWYLAGDITIYLANFTLPDSEFLKFKPEAAFLPEWVGINADLS 357  
 QY 328 PLLKALTP 335  
 Db 358 PLWTLAKP 365

RESULT 11  
 AAR70422  
 ID AAR70422 standard; Protein; 365 AA.  
 XX AC AAR70422;  
 XX DT 27-FEB-1996 (first entry)  
 XX DE 2-Alpha-fucosyltransferase.  
 XX KW GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyltransferase;  
 KW glycosyltransferase; glycosylation; oligosaccharide; glycoprotein;  
 KW glycolipid; transgenic animal; cattle; milk.  
 XX OS Homo sapiens.  
 XX PN WO9524495-Al.  
 XX PD 14-SEP-1995.  
 XX PF 24-JAN-1995; 95WO-US00967.  
 XX PR 09-MAR-1994; 94US-0208889.  
 XX PA (ABBO ) ABBOTT LAB.  
 XX PI Cummings RD, Kopchik JJ, Moreman KW, Mukerji P;  
 PI Pierce JM, Prieto PA, Smith DF;  
 XX DR WPI; 1995-328284/42.  
 XX DR N-PSDB; AAT01083.  
 XX PT New transgenic non-human mammal milk prods - contg. heterologous  
 PT components produced as secondary gene prods. of an heterologous gene  
 XX PS Example 1; Page 62-64; 83pp; English.  
 XX CC 2-Alpha-fucosyltransferase (AAR70422) is encoded by a cDNA clone  
 CC (AAT01083) isolated from a human epidermal carcinoma A431 library.  
 CC The enzyme can be expressed in the milk of a transgenic mammal,  
 CC esp. cow. This allows large-scale prodn. of oligosaccharides and

```
CC glycosylated proteins and lipids in the milk.
XX
SQ Sequence 365 AA;
Query Match 54.7%; Score 1051.5; DB 16; Length 365;
Best Local Similarity 63.3%; Pred. No. 4.3e-106;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 30 GNTSEPMRDSQHGNGELRGMTINSICRLGNQMGYATLFAARMNGRLAFIPASMH 89
DB 59 GTAMGPNASSCPOH-PASLSGTWTVYPNGRFGNQMGQVATLLALAQLNGRAFIIPAMH 117
QY 90 NALAPFTRISLVLSHDTAKKIPWQNYHLNDWMEERYHIPGHFVFTGYPGCSWTFYHHL 149
DB 118 AALAPFRTILVLAPEVDSRTPWRELQHDWMESEYADLRDPFLKLSGFGCSWTFHHL 177
QY 150 RPEILKEFTLHDHVEEAQAFRLGRV--NGSOPSTFVGHVHVRGDDYVHVMNWKGVVA 207
DB 178 REQIRREFTLHDLREEAQSVLGQLRGRTGDRPRTFVGHVHVRGDDYLVQMPQKRWGVG 237
QY 208 DRGYLEKALDMFRARYSSPVVVTSGMAWCRENINASRGDVFVFNAGNGIEGSPAKDFALL 267
DB 238 DSAYLRQAMDWFARHEAPVFPVVTSGMEWCKENIDTSQDVTFAGDGGEATPWKDFALL 297
QY 268 TCNNTIMTIGTFGWAAYLAGDIIYLANIYLPDPSFLKVPKPEAAFLPEWVGIPADLS 327
DB 298 TCNNTIMTIGTFGWAAYLAGDIIYLANIYLPDSEFLKIFKPEAAFLPEWVGINADLS 357
QY 328 PLLKALTP 335
DB 358 PLWTLAKP 365
RESULT 12
AAR70421
ID AAR70421 standard; Protein; 365 AA.
AC AAR70421;
XX
XX 27-FEB-1996 (first entry)
DE
DE 2-Alpha-fucosyltransferase.
KW GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyltransferase;
KW glycosyltransferase; oligosaccharide; glycoprotein; glycolipid;
KW transgenic animal; cattle; glycosylation; milk.
XX
XX Homo sapiens.
XX W09524488-A1.
XX 14-SEP-1995.
XX 24-JAN-1995; 95WO-US01147.
XX 09-MAR-1994; 94US-0209132.
XX (ABBO ) ABBOTT LAB.
XX Cummings RD, Kopchik JU, Moremen KW, Mukerji P;
XX Pierce JM, Prieto PA, Smith DF;
XX WPI: 1995-328279/42.
XX N-FSDB; AAT01082.
XX
XX Transgenic animal expressing heterologous catalyst - used in
XX metabolite prodn esp. glycosyltransferase for prodn. of
XX oligosaccharide(s) and glyco-conjugate(s) in the milk, useful in
XX foods, pharmaceuticals, etc.
XX
XX Example 1; Page 62-64; 84pp; English.
XX
XX Alpha-1,2-fucosyltransferase (AAR70421) is encoded by a cDNA clone
CC
```

```
CC (AAT01082) isolated from a human epidermal carcinoma cell line A431
XX cDNA library. The enzyme can be expressed in the milk of a
XX transgenic mammal, esp. cow. This allows large-scale prodn. of
XX oligosaccharides or glycosylated proteins and lipids in the milk.
XX
SQ Sequence 365 AA;
Query Match 54.7%; Score 1051.5; DB 16; Length 365;
Best Local Similarity 63.3%; Pred. No. 4.3e-106;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 30 GNTSEPMRDSQHGNGELRGMTINSICRLGNQMGYATLFAARMNGRLAFIPASMH 89
DB 59 GTAMGPNASSCPOH-PASLSGTWTVYPNGRFGNQMGQVATLLALAQLNGRAFIIPAMH 117
QY 90 NALAPFTRISLVLSHDTAKKIPWQNYHLNDWMEERYHIPGHFVFTGYPGCSWTFYHHL 149
DB 118 AALAPFRTILVLAPEVDSRTPWRELQHDWMESEYADLRDPFLKLSGFGCSWTFHHL 177
QY 150 RPEILKEFTLHDHVEEAQAFRLGRV--NGSOPSTFVGHVHVRGDDYVHVMNWKGVVA 207
DB 178 REQIRREFTLHDLREEAQSVLGQLRGRTGDRPRTFVGHVHVRGDDYLVQMPQKRWGVG 237
QY 208 DRGYLEKALDMFRARYSSPVVVTSGMAWCRENINASRGDVFVFNAGNGIEGSPAKDFALL 267
DB 238 DSAYLRQAMDWFARHEAPVFPVVTSGMEWCKENIDTSQDVTFAGDGGEATPWKDFALL 297
QY 268 TCNNTIMTIGTFGWAAYLAGDIIYLANIYLPDPSFLKVPKPEAAFLPEWVGIPADLS 327
DB 298 TCNNTIMTIGTFGWAAYLAGDIIYLANIYLPDSEFLKIFKPEAAFLPEWVGINADLS 357
QY 328 PLLKALTP 335
DB 358 PLWTLAKP 365
RESULT 13
AAR90572
ID AAR90572 standard; Protein; 365 AA.
AC AAR90572;
XX
XX 08-APR-1996 (first entry)
DE
DE Human H-transferase.
KW H-transferase; xenograft hyperacute rejection; transplantation;
KW glycosyltransferase; galactose alpha(1,3) galactose.
XX
XX Homo sapiens.
XX W09534202-A1.
XX 21-DEC-1995.
XX 14-JUN-1995; 95WO-US07554.
XX 21-JUL-1994; 94US-0278282.
XX 15-JUN-1994; 94US-0260201.
XX (ALEX-) ALEXION PHARM INC.
XX (AUST-) AUSTIN RES INST.
XX Fodor WL, McKenzie IFC, Rother RP, Sandrin MS, Squinto SP;
XX WPI: 1996-049326/05.
XX N-FSDB; AAT12238.
XX
XX Redn. of rejection of xenogeneic cells following transplantation
XX by introducing a vector expressing fucosyltransferase into the
XX cells
XX
XX Example 1; Page 45-47; 69pp; English.
XX
XX
```

```
XX The human H-transferase (AA090572) product of a cDNA clone (AA012238);
CC pred. from human epidermoid carcinoma cells (ATCC CRL 155 A-431)
CC can be expressed in xenogeneic organs, tissues and cells using
CC a vector such as PAPX-1 (AA012238). This results in decreased
CC expression of the non-human antigen galactose alpha(1,3) galactose
CC on the surface of the organs etc. so that hyperacute rejection is
CC reduced upon transplantation to humans.
XX
SQ Sequence 365 AA;
Query Match 54.7%; Score 1051.5; DB 17; Length 365;
Best Local Similarity 63.3%; Pred. No. 4.3e-106;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 30 GNTPEPMRRDSECHNGELRGMTFINSIGRLGNQMGVATLFLALRMNGRLAFIPASMH 89
Db 59 GTAMGNASSCCQOH-PASLSGWTVPYNGRFGNQGYATLLALQNLGRRAFILPANH 117
QY 90 NALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEERYHLPFGHFVFTGYPCSWTFYHHL 149
Db 118 AALAPVFRITLPVLAPEVDSRTPWRELQLDHWMSEYADLRDPFLKLSGFPCCSWTFEHL 177
QY 150 RPEILKEFTLHDHVRREAQAFRLGLRV--NGSOPSTFVGHVHVRGDYVHVMPNWKGVVA 207
Db 178 REQIRREFTLHDHLRBEAQSVLGQLRGTRGDRPTFVGHVHVRGDYLVQVMPQWKGVVG 237
QY 208 DRGYLEKALDMFRARYSSPVFVVTSGMAWCRNINASRGDVVFAGNGIEGSPAKDFALL 267
Db 238 DSAYLQAMDWFRARHEAPVFVVTSGMEWCKENIDTSQGDVTFAGDQGEATPWKDFALL 297
QY 268 TCNHTIMTIGTFGWAAYLAGDGTIYLANFTLPDSEFLKIEKPEAAFLPEWVGIPADLS 327
Db 298 TCNHTIMTIGTFGWAAYLAGDGTIYLANFTLPDSEFLKIEKPEAAFLPEWVGIPADLS 357
QY 328 PLLKALTP 335
Db 358 PLWTIAKP 365
RESULT 14
AAW23805
ID AAW23805 standard; Protein; 365 AA.
XX
AC AAW23805;
XX
DT 15-SEP-1997 (first entry)
XX
DE Human alpha 1,2 fucosyltransferase (EC 2.4.1.69).
XX
KW Human GDP-L-fucose:beta-D-galactoside-2-alpha-L-fucosyltransferase;
KW alpha 1,2 FT; transgene; transgenic mouse; animal model;
KW intestinal adhesion; Helicobacter pylori infection; stomach;
KW small intestine; gut; epithelial cell; surface receptor;
KW carbohydrate antigen; gastritis; peptic ulcer; neoplasia;
KW gastric adenocarcinoma; Lewis antigen; fucosylation.
XX
OS Homo sapiens.
XX
PN US5625124-A.
XX
PD 29-APR-1997.
XX
PF 11-JUL-1994; 94US-0273411.
XX
PR 11-JUL-1994; 94US-0273411.
PA (UNITW ) UNIV WASHINGTON.
XX
PI Falk P, Gordon JI;
XX
DR WPI: 1997-258275/23.
DR N-PSDB; AAI76768.
```

```
XX Animal model for Helicobacter pylori infection - comprising
PT transgenic mouse expressing human enzyme promoting intestinal
PT adhesion
XX
PS Example 1; Columns 13-16; 24pp; English.
XX
CC A claimed transgenic mouse expresses, in its intestinal epithelial
CC cells, the enzyme human GDP-L-fucose:beta-D-galactoside 2-alpha-L-
CC fucosyltransferase (also called alpha 1,2 FT) or human GDP-L-fucose:
CC beta-D-N-acetylglucosamide 3/4-alpha-L-fucosyltransferase (also
CC called alpha 1,3/4 FT). The enzyme is expressed under the
CC control of a gut epithelial cell-specific promoter and Helicobacter
CC pylori adheres to the transgenic cells. The transgenic mouse and
CC intestinal epithelial cells from it are useful as models for screening
CC compounds for the ability to inhibit adhesion of H. pylori to gut
CC epithelial cells. The present sequence represents human alpha 1,2 FT
CC and was published by Larsen et al., Proc.Natl.Acad.Sci.USA, 87:
CC 6674-6678 (1990).
XX
SQ Sequence 365 AA;
Query Match 54.7%; Score 1051.5; DB 18; Length 365;
Best Local Similarity 63.3%; Pred. No. 4.3e-106;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 30 GNTPEPMRRDSECHNGELRGMTFINSIGRLGNQMGVATLFLALRMNGRLAFIPASMH 89
Db 59 GTAMGNASSCCQOH-PASLSGWTVPYNGRFGNQGYATLLALQNLGRRAFILPANH 117
QY 90 NALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEERYHLPFGHFVFTGYPCSWTFYHHL 149
Db 118 AALAPVFRITLPVLAPEVDSRTPWRELQLDHWMSEYADLRDPFLKLSGFPCCSWTFEHL 177
QY 150 RPEILKEFTLHDHVRREAQAFRLGLRV--NGSOPSTFVGHVHVRGDYVHVMPNWKGVVA 207
Db 178 REQIRREFTLHDHLRBEAQSVLGQLRGTRGDRPTFVGHVHVRGDYLVQVMPQWKGVVG 237
QY 208 DRGYLEKALDMFRARYSSPVFVVTSGMAWCRNINASRGDVVFAGNGIEGSPAKDFALL 267
Db 238 DSAYLQAMDWFRARHEAPVFVVTSGMEWCKENIDTSQGDVTFAGDQGEATPWKDFALL 297
QY 268 TCNHTIMTIGTFGWAAYLAGDGTIYLANFTLPDSEFLKIEKPEAAFLPEWVGIPADLS 327
Db 298 TCNHTIMTIGTFGWAAYLAGDGTIYLANFTLPDSEFLKIEKPEAAFLPEWVGIPADLS 357
QY 328 PLLKALTP 335
Db 358 PLWTIAKP 365
RESULT 15
AAW13640
ID AAW13640 standard; Protein; 365 AA.
XX
AC AAW13640;
XX
DT 19-JUN-1997 (first entry)
XX
DE Human alpha(1,2)-fucosyltransferase.
XX
KW Alpha(1,2)-fucosyltransferase; glycosylation; oligosaccharide;
KW blood group H.
XX
OS Homo sapiens.
XX
PN WO9709421-A1.
XX
PD 13-MAR-1997.
XX
PF 06-SEP-1996; 96WO-US13616.
XX
PR 08-SEP-1995; 95US-0525058.
```



GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:03:22 ; Search time 14.7192 Seconds  
(without alignments)  
715.625 Million cell updates/sec

Title: US-10-040-863-8\_COPY\_23\_380

Perfect score: 1921

Sequence: 1 STIIHQQRIYVKQLPLSEKE.....RSHFHLKAKGVTCYVAGRAF 358

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PTCUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1921	100.0	390	4	US-09-298-886-8
2	1896	98.7	353	4	US-09-298-886-10
3	1451	75.5	340	4	US-09-254-077A-6
4	1451	75.5	340	4	US-09-254-077A-7
5	1407.5	73.3	343	1	US-08-395-800A-8
6	1407.5	73.3	343	4	US-09-254-077A-8
7	1397	72.7	344	4	US-09-298-886-11
8	1396.5	72.7	347	4	US-09-254-077A-9
9	1086	56.5	373	4	US-09-254-077A-12
10	1051.5	54.7	333	5	PCT-US91-00899-11
11	1051.5	54.7	365	1	US-07-914-281-6
12	1051.5	54.7	365	1	US-08-393-246-6
13	1051.5	54.7	365	1	US-08-273-411-1
14	1051.5	54.7	365	1	US-08-525-058A-6
15	1051.5	54.7	365	1	US-08-395-800A-6
16	1051.5	54.7	365	1	US-08-395-800A-10
17	1051.5	54.7	365	2	US-08-696-731-6
18	1051.5	54.7	365	4	US-09-042-531-6
19	1051.5	54.7	365	4	US-09-254-077A-11
20	1051.5	54.7	365	5	PCT-US91-00899-12
21	1046.5	54.5	365	4	US-09-151-592-2
22	1046.5	54.5	365	4	US-09-254-077A-10
23	696.5	36.3	222	1	US-08-395-800A-2
24	439	22.9	102	1	US-08-395-800A-3
25	141	7.3	300	4	US-09-433-598-2
26	97	5.0	183	1	US-08-167-035-33
27	97	5.0	183	1	US-08-208-887A-33

28	97	5.0	183	2	US-08-539-005-33	Sequence 33, Appl
29	97	5.0	183	4	US-09-280-598-35	Sequence 35, Appl
30	96.5	5.0	355	4	US-09-390-131-5	Sequence 5, Appl
31	91.5	4.8	801	1	US-07-906-349A-6	Sequence 6, Appl
32	91	4.7	442	1	US-08-220-151-22	Sequence 22, Appl
33	91	4.7	442	1	US-08-413-118-22	Sequence 22, Appl
34	91	4.7	442	3	US-08-473-446-22	Sequence 22, Appl
35	91	4.7	442	6	5470718-3	Patent No. 5470718
36	87	4.5	217	1	US-08-167-035-6	Sequence 6, Appl
37	87	4.5	217	1	US-08-208-887A-6	Sequence 6, Appl
38	87	4.5	217	2	US-08-539-005-6	Sequence 6, Appl
39	87	4.5	217	2	US-08-815-176-3	Sequence 3, Appl
40	87	4.5	217	2	US-08-815-176-4	Sequence 4, Appl
41	87	4.5	217	4	US-08-664-962B-6	Sequence 6, Appl
42	87	4.5	217	4	US-09-311-743-6	Sequence 6, Appl
43	87	4.5	217	4	US-09-280-598-6	Sequence 6, Appl
44	87	4.5	217	4	US-09-197-344-3	Sequence 3, Appl
45	87	4.5	217	4	US-09-197-344-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-09-298-886-8  
; Sequence 8, Application US/09298886  
; Patent No. 5329170  
; GENERAL INFORMATION:  
; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
; FILE REFERENCE: 8511-029  
; CURRENT APPLICATION NUMBER: US/09/298,886  
; CURRENT FILING DATE: 1999-04-26  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 380  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-298-886-8

Query Match 100.0%; Score 1921; DB 4; Length 380;  
Best Local Similarity 100.0%; Pred. No. 8.7e-214;  
Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	STIIHQQRIYVKQLPSEKELPMTQMSNGNTSEPMRRDSEQHNGELRGMTINSIGR	60
Db	23	STIIHQQRIYVKQLPSEKELPMTQMSNGNTSEPMRRDSEQHNGELRGMTINSIGR	82
QY	61	LGNGMGVATLFAIARNGRLAFTPSMHNALAFRTSLPVLHSDTAKKIPWONYHLND	120
Db	83	LGNGMGVATLFAIARNGRLAFTPSMHNALAFRTSLPVLHSDTAKKIPWONYHLND	142
QY	121	WMERYRHPGHVFRTGYPCSWTFYHHLRPEILKEFTLHDHVEEAQAFRLGLRVNGSQ	180
Db	143	WMERYRHPGHVFRTGYPCSWTFYHHLRPEILKEFTLHDHVEEAQAFRLGLRVNGSQ	202
QY	181	PSTFVGTVHVRGDDVHVHNPVNVKGVADRGYLEKALDMFRARYSSPVFVTSNGMACRE	240
Db	203	PSTFVGTVHVRGDDVHVHNPVNVKGVADRGYLEKALDMFRARYSSPVFVTSNGMACRE	262
QY	241	NINASRGDVFAGNGIEGSPAKDFALLTQCNTTMTCTGFIWAAYLAGGDTIYLANVTL	300
Db	263	NINASRGDVFAGNGIEGSPAKDFALLTQCNTTMTCTGFIWAAYLAGGDTIYLANVTL	322
QY	301	PDSPFLKVFPEAAFLPEWVGIPADLSPLKALTPACPRSHFHLKAKGVTCYVAGRAF	358
Db	323	PDSPFLKVFPEAAFLPEWVGIPADLSPLKALTPACPRSHFHLKAKGVTCYVAGRAF	380

RESULT 2

```
US-09-298-886-10
; Sequence 10, Application US/09298886
; Patent No. 6329170
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUOSYLTRANSFERASE AND USES
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/298,886
; CURRENT FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-298-886-10

Query Match      98.7%; Score 1896; DB 4; Length 353;
Best Local Similarity 100.0%; Pred. No. 6.le-211;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LQORIVKLOPLSEKELPMTQSSNTESPEMRDSEQHNGELRGMFTINSIGRLGNOM 65
DQ 1 LQORIVKLOPLSEKELPMTQSSNTESPEMRDSEQHNGELRGMFTINSIGRLGNOM 60
QY 66 GEYATLALARMNGRLAFIPASMHNALAPIRISLPVLHSDTAKKIPWNYHLNDWMEER 125
DQ 61 GEYATLALARMNGRLAFIPASMHNALAPIRISLPVLHSDTAKKIPWNYHLNDWMEER 120
QY 126 YRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFGLRUVNGSQPSTEV 185
DQ 121 YRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFGLRUVNGSQPSTEV 180
QY 186 GYHVRGDDYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMWCARENINAS 245
DQ 181 GYHVRGDDYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMWCARENINAS 240
QY 246 RGDVVFAGNGIEGSPAKDFALLTQCNHTIMTGTGFWAAAYLAGGDTIYLANNTLPDSPF 305
DQ 241 RGDVVFAGNGIEGSPAKDFALLTQCNHTIMTGTGFWAAAYLAGGDTIYLANNTLPDSPF 300
QY 306 LKVPKPEAFLEWVGIPADLSPLKALTPACPRSHFHLKAKGVTCYVAGRAF 358
DQ 301 LKVPKPEAFLEWVGIPADLSPLKALTPACPRSHFHLKAKGVTCYVAGRAF 353

RESULT 3
US-09-254-077A-6
; Sequence 6, Application US/09254077A
; Patent No. 6399758
; GENERAL INFORMATION:
; APPLICANT: SANDRIN, MAURO S. F.
; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES
; FILE REFERENCE: 30562.5USWO
; CURRENT APPLICATION NUMBER: US/09/254,077A
; CURRENT FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: PCT/AU97/00540
; PRIOR FILING DATE: 1997-08-22
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-254-077A-6

Query Match      75.5%; Score 1451; DB 4; Length 340;
Best Local Similarity 81.9%; Pred. No. 2e-159;
Matches 340; Conservative 26; Mismatches 20; Indels 14; Gaps 3;

QY 1 STIIHQORIVKLOPLSEKELPMTQSSNTESPEMRDSEQHNGELRGMFTINSIGR 60
DQ 23 STIIHQORIVKLOPLSEKELPMTQSSNTESPEMRDSEQHNGELRGMFTINSIGR 58
QY 61 LGNOMGEYATLALARMNGRLAFIPASMHNALAPIRISLPVLHSDTAKKIPWNYHLND 120
DQ 69 LGNOMGEYATLALARMNGRLAFIPASMHNALAPIRISLPVLHSDTAKKIPWNYHLND 128
QY 121 WMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFGLRUVNGSQ 180
DQ 129 WMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFGLRUVNGSR 188
QY 181 PSTFVGHVRRGDYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMWCARE 240
DQ 189 PSTFVGHVRRGDYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMWCARE 248
QY 241 NINASRGDVVFAGNGIEGSPAKDFALLTQCNHTIMTGTGFWAAAYLAGGDTIYLANNTL 300
DQ 249 NINASRGDVVFAGNGIEGSPAKDFALLTQCNHTIMTGTGFWAAAYLAGGDTIYLANNTL 308

Query Match      75.5%; Score 1451; DB 4; Length 340;
Best Local Similarity 81.9%; Pred. No. 2e-159;
Matches 340; Conservative 26; Mismatches 20; Indels 14; Gaps 3;

QY 1 STIIHQORIVKLOPLSEKELPMTQSSNTESPEMRDSEQHNGELRGMFTINSIGR 60
DQ 23 STIIHQORIVKLOPLSEKELPMTQSSNTESPEMRDSEQHNGELRGMFTINSIGR 58
QY 61 LGNOMGEYATLALARMNGRLAFIPASMHNALAPIRISLPVLHSDTAKKIPWNYHLND 120
DQ 69 LGNOMGEYATLALARMNGRLAFIPASMHNALAPIRISLPVLHSDTAKKIPWNYHLND 128
QY 121 WMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFGLRUVNGSQ 180
DQ 129 WMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFGLRUVNGSR 188
QY 181 PSTFVGHVRRGDYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMWCARE 240
DQ 189 PSTFVGHVRRGDYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMWCARE 248
QY 241 NINASRGDVVFAGNGIEGSPAKDFALLTQCNHTIMTGTGFWAAAYLAGGDTIYLANNTL 300
DQ 249 NINASRGDVVFAGNGIEGSPAKDFALLTQCNHTIMTGTGFWAAAYLAGGDTIYLANNTL 308
```

QY 301 PDSPFLKVKPEAAFLPEWVGIPADLSPLLK 331  
Db 309 PDSPFLKVKPEAAFLPEWVGIPADLSPLLK 339

RESULT 5  
US-08-395-800A-8

; Sequence 8, Application US/08395800A  
; Patent No. 5807732

; GENERAL INFORMATION:

; APPLICANT: LOWE, JOHN B

; APPLICANT: LENNON, GREGORY

; APPLICANT: ROQUIER, SYLVIE

; APPLICANT: GIORGI, DOMINIQUE

; APPLICANT: KELLY, ROBERT J

; TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE

; TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE

; TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF

; TITLE OF INVENTION: GENOTYPING A PERSON

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBION, SPIVAK, MCLELLAND, MAIER & NEUSTADT

; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400

; CITY: ARLINGTON

; STATE: VIRGINIA

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/395,800A

; FILING DATE: 28-FEB-1995

; CLASSIFICATION: 435

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 413-3000

; TELEFAX: (703) 413-2220

; TELEX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 343 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-395-800A-8

Query Match 73.3%; Score 1407.5; DB 1; Length 343;

Best Local Similarity 79.2%; Pred. No. 2.2e-154;

Matches 262; Conservative 26; Mismatches 32; Indels 11; Gaps 2;

QY 1 STIIHQQRIVKQLPSEKELPMTOMSSGNTSEPMRRDSEQHNGELRGMTINSIGR 60

Db 23 STIFHVQQRLLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGMTINAIGR 71

QY 61 LGNOMGEYATLFAALRMNGRLAFIPASMHNALAFIRISLPVLHSDTAKKIPQNYHLND 120

Db 72 LGNOMGEYATLYALAKMNGRPAFIPAQMHSHTLAFIPRITLPVLHSATASRIPQNYHLND 131

QY 121 WMEERYRHIPGHEVVRTGYPCSWTFYHHLRQELQETLHDHVREAAQALRLGRLVNGSQ 180

Db 132 WMEERYRHIPGEYVRTGYPCSWTFYHHLRQELQETLHDHVREAAQALRLGRLVNGSR 191

QY 181 PSTFVGHVVRGDIYVHVMKGVVADRGLYLEKALDMFRARYSSPVVYTSNGMAWCRE 240

Db 192 PGTFVGHVVRGDIYVHVMKGVVADRGLYLOALDMFRARYSSLIFFVYTSNGMAWCRE 251

QY 241 NINASRCDVVFAGNGIEGSPAKDFALLTQCNHMTIGTGTGIAWAAYLAGDITIYLANYYL 300

Db 252 NIDTSHGDVVFAGDIEGSPAKDFALLTQCNHMTIGTGTGIAWAAYLTGGDITIYLANYYL 311

QY 301 PDSPFLKVKPEAAFLPEWVGIPADLSPLLK 331  
Db 312 PDSPFLKVKPEAAFLPEWVGIPADLSPLLK 342

RESULT 6

US-09-254-077A-8

; Sequence 8, Application US/09254077A

; Patent No. 6399758

; GENERAL INFORMATION:

; APPLICANT: SANDRIN, MAURO S.

; APPLICANT: MCKENZIE, IAN C. F.

; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES

; FILE REFERENCE: 30562.5USMO

; CURRENT APPLICATION NUMBER: US/09/254,077A

; CURRENT FILING DATE: 1999-06-11

; PRIOR APPLICATION NUMBER: PCT/AU97/00540

; PRIOR FILING DATE: 1997-08-22

; PRIOR APPLICATION NUMBER: PO 1823

; PRIOR FILING DATE: 1996-08-23

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 8

; LENGTH: 343

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-254-077A-8

Query Match 73.3%; Score 1407.5; DB 4; Length 343;

Best Local Similarity 79.2%; Pred. No. 2.2e-154;

Matches 262; Conservative 26; Mismatches 32; Indels 11; Gaps 2;

QY 1 STIIHQQRIVKQLPSEKELPMTOMSSGNTSEPMRRDSEQHNGELRGMTINSIGR 60

Db 23 STIFHVQQRLLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGMTINAIGR 71

QY 61 LGNOMGEYATLFAALRMNGRLAFIPASMHNALAFIRISLPVLHSDTAKKIPQNYHLND 120

Db 72 LGNOMGEYATLYALAKMNGRPAFIPAQMHSHTLAFIPRITLPVLHSATASRIPQNYHLND 131

QY 121 WMEERYRHIPGHEVVRTGYPCSWTFYHHLRQELQETLHDHVREAAQALRLGRLVNGSQ 180

Db 132 WMEERYRHIPGEYVRTGYPCSWTFYHHLRQELQETLHDHVREAAQALRLGRLVNGSR 191

QY 181 PSTFVGHVVRGDIYVHVMKGVVADRGLYLEKALDMFRARYSSPVVYTSNGMAWCRE 240

Db 192 PGTFVGHVVRGDIYVHVMKGVVADRGLYLOALDMFRARYSSLIFFVYTSNGMAWCRE 251

QY 241 NINASRCDVVFAGNGIEGSPAKDFALLTQCNHMTIGTGTGIAWAAYLAGDITIYLANYYL 300

Db 252 NIDTSHGDVVFAGDIEGSPAKDFALLTQCNHMTIGTGTGIAWAAYLTGGDITIYLANYYL 311

QY 301 PDSPFLKVKPEAAFLPEWVGIPADLSPLLK 331

Db 312 PDSPFLKVKPEAAFLPEWVGIPADLSPLLK 342

RESULT 7

US-09-298-886-11

; Sequence 11, Application US/09298886

; Patent No. 6329170

; GENERAL INFORMATION:

; APPLICANT: Eric H. Holmes et al.

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE

; TITLE OF INVENTION: GM1-SPECIFIC ALPHA-2 FUCOSYLTRANSFERASE AND USES

; FILE REFERENCE: 8511-C29

; CURRENT APPLICATION NUMBER: US/09/298,886

; CURRENT FILING DATE: 1999-04-26

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 11

; LENGTH: 344







;; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
;; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
;; TITLE OF INVENTION: GLYCOPOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
;; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
;; NUMBER OF SEQUENCES: 14  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: OBLON, SPIVAK, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
;; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
;; CITY: Arlington  
;; STATE: Virginia  
;; COUNTRY: U.S.A.  
;; ZIP: 22202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/393,246  
;; FILING DATE:  
;; CLASSIFICATION: 530  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/220,433  
;; FILING DATE: 30-MAR-1994  
;; APPLICATION NUMBER: US 07/914,281  
;; FILING DATE: 20-JUL-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Lavalleye, Jean-Paul M. P.  
;; REGISTRATION NUMBER: 31,451  
;; REFERENCE/DOCKET NUMBER: 2363-060-55  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703)521-4500  
;; TELEFAX: (703)486-2347  
;; TELEX: 248855 OPAT UR  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 365 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: protein  
US-08-393-246-6

Query Match 54.7%; Score 1051.5; DB 1; Length 365;  
Best Local Similarity 63.3%; Pred. No. 4.3e-113;  
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;  
QY 30 GNTESPMRDRDSQHNGELRGMTINSIGRLGNOMGEYATLFLALARMNGRLAFIPASMH 89  
Db 59 GTAMGNASSCCPQH-PASLSGTWTVPNGRFGNQGYATLLAQLNGRRRAIFLPAH 117  
QY 90 NALAFIRISLPVLHSDTAKKIPQWYHLNDWMEERYRHIPGHVFTGYPCSWTFVYHLL 149  
Db 118 AALAPVRITLPVLAPVDSRIPTWRELQHDWMESEYADLRDPFLKLSGPGCSWTFVHLL 177  
QY 150 RPELKEFTLHDHVREBAQAFRLRLV--NGSQPTFVGHVHVRGDYVHYMPNWKGVWA 207  
Db 178 REGIRREFTLHDHLREBAQSVLGQLRLGRGTDRPTFVGHVHVRGDYVHYMPQWRKGVWG 237  
QY 208 DRGYLEKALDMFRARYSSPVFVVTNSGMACRENINASRGDVVPFAGNIGSGSAKDFALL 267  
Db 238 DSAYLRQAMDFWARHEAPYFVVTNSGMWCKENIDTSGQDVTFAGDQGQATPWKDFALL 297  
QY 268 TQCNTHTMTIGTGFWNAAYLAGDITIYLANITLPDPSFLKVPKEAFLPEWVGIPADLS 327  
Db 298 TQCNTHTMTIGTGFWNAAYLAGDITVILANFTLPDSEFLKIFPEAFLPEWVGINADLS 357  
QY 328 PLLKALTP 335  
Db 358 PLWTAKP 365  
RESULT 13

US-08-273-411-1  
;; Sequence 1, Application US/08273411  
;; Patent No. 5625124  
;; GENERAL INFORMATION:  
;; APPLICANT: Falk, Per  
;; APPLICANT: Gordon, Jeffrey I.  
;; TITLE OF INVENTION: Animal Model for Gastro-Intestinal  
;; TITLE OF INVENTION: Disease  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Patrea L. Pabst  
;; STREET: 1100 Peachtree Street, Suite 2800  
;; CITY: Atlanta  
;; STATE: Georgia  
;; COUNTRY: USA  
;; ZIP: 30309-4530  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/273,411  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Pabst, Patrea L.  
;; REGISTRATION NUMBER: 31,284  
;; REFERENCE/DOCKET NUMBER: WU106  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (404) 815-6508  
;; TELEFAX: (404) 815-6555  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 365 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOPHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: 1..365  
;; OTHER INFORMATION: /note= "GDP-L-fucose:beta-D-Galactoside-2-alpha-L-fucosyl  
;; PUBLICATION INFORMATION:  
;; AUTHORS: Larsen, et al.  
;; JOURNAL: Proc. Nat'l Acad. Sci. USA  
;; VOLUME: 87  
;; PAGES: 6674-6678  
;; DATE: 1990  
;; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 365  
US-08-273-411-1

Query Match 54.7%; Score 1051.5; DB 1; Length 365;  
Best Local Similarity 63.3%; Pred. No. 4.3e-113;  
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;  
QY 30 GNTESPMRDRDSQHNGELRGMTINSIGRLGNOMGEYATLFLALARMNGRLAFIPASMH 89  
Db 59 GTAMGNASSCCPQH-PASLSGTWTVPNGRFGNQGYATLLAQLNGRRRAIFLPAH 117  
QY 90 NALAFIRISLPVLHSDTAKKIPQWYHLNDWMEERYRHIPGHVFTGYPCSWTFVYHLL 149  
Db 118 AALAPVRITLPVLAPVDSRIPTWRELQHDWMESEYADLRDPFLKLSGPGCSWTFVHLL 177  
QY 150 RPELKEFTLHDHVREBAQAFRLRLV--NGSQPTFVGHVHVRGDYVHYMPNWKGVWA 207  
Db 178 REGIRREFTLHDHLREBAQSVLGQLRLGRGTDRPTFVGHVHVRGDYVHYMPQWRKGVWG 237  
QY 208 DRGYLEKALDMFRARYSSPVFVVTNSGMACRENINASRGDVVPFAGNIGSGSAKDFALL 267

Db 238 DSAVLQAMDFRAREHAPVFTVTSNGMCKENIDTSQGDVTFAGDQGEATPKWKDFALL 297  
QY 266 TQCNHTMTTGTGCIWAAYLAGGDTIYLANVTLPDPSFLKVKFPEAAFLPEWVGIPADLS 327  
Db 298 TQCNHTMTTGTGCIWAAYLAGGDTIYLANVTLPDPSFLKVKFPEAAFLPEWVGIPADLS 357  
QY 328 PLLKALTP 335  
Db 358 PLWTLAKP 365

RESULT 14  
US-08-525-058A-6  
; Sequence 6, Application US/08525058A  
; Patent No. 5770420  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/525,058A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye, Jean-Paul M. P.  
; REGISTRATION NUMBER: 31,451  
; REFERENCE/DOCKET NUMBER: 2363-060-55  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)521-4500  
; TELEFAX: (703)486-2347  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 365 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-525-058A-6

Query Match 54.7%; Score 1051.5; DB 1; Length 365;  
Best Local Similarity 63.3%; Pred. No. 4.3e-113;  
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

QY 30 GNTESPWRDSEOHGNGELRGMTINSIGRLGNQMGYATLFLALARNGLAFIPASMH 89  
Db 59 GTAMGPNASSCPOH-PASISGTTWTVPNGRFGNQMGYATLLALALQNGRRRAFIIPAMH 117  
QY 90 NALAPIRISLPVLHSDTAKIPWNYHLNDWMEERYRHIPGHVFTGYPCSWTFYHHL 149  
Db 118 AALAPVFRITLPVLAPEVDSKTPWRELQLDHDMSEEVADLRDPFLKLSGFPSCWTFYHHL 177  
QY 150 RPEILKEFTLHDHVRERAAQAFRLGLRV--NGSQPSTFGVHVRRGDYVHYMPNWKGVVA 207  
Db 178 REQIRREFTLHDHLEERAAQSVLGLRLGRTGDRPTFGVHVRRGDYLVQMPQWKGVVG 237  
QY 208 DRGYLEKALDMFRARYSPVFTVTSNGMCKENINASRGDVFVAGNGIEGSPAKDFALL 267  
Db 238 DSAVLQAMDFRAREHAPVFTVTSNGMCKENIDTSQGDVTFAGDQGEATPKWKDFALL 297

QY 268 TQCNHTMTTGTGCIWAAYLAGGDTIYLANVTLPDPSFLKVKFPEAAFLPEWVGIPADLS 327  
Db 298 TQCNHTMTTGTGCIWAAYLAGGDTIYLANVTLPDPSFLKVKFPEAAFLPEWVGIPADLS 357  
QY 328 PLLKALTP 335  
Db 358 PLWTLAKP 365

RESULT 15  
US-08-395-800A-6  
; Sequence 6, Application US/08395800A  
; Patent No. 5807732  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; APPLICANT: LENNON, GREGORY  
; APPLICANT: ROQUIER, SYLVIE  
; APPLICANT: GIORGI, DOMINIQUE  
; APPLICANT: KELLY, ROBERT J.  
; TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE  
; TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE  
; TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF  
; TITLE OF INVENTION: GENOTYPING A PERSON  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/395,800A  
; FILING DATE: 28-FEB-1995  
; CLASSIFICATION: 435  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 365 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-395-800A-6

Query Match 54.7%; Score 1051.5; DB 1; Length 365;  
Best Local Similarity 63.3%; Pred. No. 4.3e-113;  
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

QY 30 GNTESPWRDSEOHGNGELRGMTINSIGRLGNQMGYATLFLALARNGLAFIPASMH 89  
Db 59 GTAMGPNASSCPOH-PASISGTTWTVPNGRFGNQMGYATLLALALQNGRRRAFIIPAMH 117  
QY 90 NALAPIRISLPVLHSDTAKIPWNYHLNDWMEERYRHIPGHVFTGYPCSWTFYHHL 149  
Db 118 AALAPVFRITLPVLAPEVDSKTPWRELQLDHDMSEEVADLRDPFLKLSGFPSCWTFYHHL 177  
QY 150 RPEILKEFTLHDHVRERAAQAFRLGLRV--NGSQPSTFGVHVRRGDYVHYMPNWKGVVA 207  
Db 178 REQIRREFTLHDHLEERAAQSVLGLRLGRTGDRPTFGVHVRRGDYLVQMPQWKGVVG 237  
QY 208 DRGYLEKALDMFRARYSPVFTVTSNGMCKENINASRGDVFVAGNGIEGSPAKDFALL 267  
Db 238 DSAVLQAMDFRAREHAPVFTVTSNGMCKENIDTSQGDVTFAGDQGEATPKWKDFALL 297

QY 268 TOCNHTIMTIGTFGIWAAYLAGDITIYLANVTLPDSFPLKVKPEAAFLPEWVGIPADLS 327  
Db 298 TOCNHTIMTIGTFGEFWAAYLAGDITIYLANVTLPDSFPLKVKPEAAFLPEWVGINADLS 357  
QY 328 PILKALTP 335  
Db 358 PWTIAKP 365

Search completed: May 27, 2003, 15:11:51  
Job time : 15.7192 secs

[illegible]

2058	9	US-09-815-379-17	
217	10	US-09-765-298A-6	
738	9	US-10-235-521-1	
344	9	US-09-896-813A-56	
344	10	US-09-815-242-11366	
747	9	US-10-028-072-212	
747	9	US-10-131-049-212	
747	9	US-10-123-904-212	
747	9	US-10-140-470-212	
747	9	US-10-175-746-212	
747	9	US-10-176-918-212	
747	9	US-10-176-921-212	
747	9	US-10-137-865-212	
747	9	US-10-140-474-212	
747	9	US-10-142-431-212	
747	9	US-10-143-114-212	
747	9	US-10-140-002-212	
747	9	US-10-142-419-212	
747	9	US-10-123-262-212	
747	9	US-10-123-423-212	
747	9	US-10-121-050-212	
747	9	US-10-141-755-212	
747	9	US-10-143-032-212	
747	9	US-10-123-108-212	
747	9	US-10-123-236-212	
747	9	US-10-123-261-212	

20	89	4.6
21	87	4.5
22	85.5	4.5
23	84.5	4.4
24	84	4.4
25	83.5	4.3
26	83.5	4.3
27	83.5	4.3
28	83.5	4.3
29	83.5	4.3
30	83.5	4.3
31	83.5	4.3
32	83.5	4.3
33	83.5	4.3
34	83.5	4.3
35	83.5	4.3
36	83.5	4.3
37	83.5	4.3
38	83.5	4.3
39	83.5	4.3
40	83.5	4.3
41	83.5	4.3
42	83.5	4.3
43	83.5	4.3
44	83.5	4.3
45	83.5	4.3

\_\_\_\_\_

3

... ..

AGH

593

78 d. 46 .ic AKO 37

5—  
n

4  
pu  
ti  
rou  
36  
RS  
SS

5 C  
RC  
w/1  
29  
...  
ues  
net

si  
00  
de  
s  
38  
...  
si  
an

29

CO  
19  
ng  
:0  
CO  
PL  
ex  
72  
ch  
10  
e

3 - VK G 2

0-0-QQ

py  
S  
27  
0-  
IR  
UM  
P  
93  
S  
h:  
h:

te	Ma	US	19	BL	Ga	37	hi	en	en
			1						

d

in  
so  
:  
ta  
:  
mb  
DE  
DE

ro  
on  
e:  
ec  
en  
in  
ch  
l  
mu  
mu

COM  
Run  
Time  
Pe  
Se  
SC  
Se  
To  
Mil  
Ma

## ALIGNMENTS

```

RESULT 1
US-09-999-672-8
: Sequence 8, Application US/09999672
: Patent No. US20020127655A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Eric H. Holmes et al.
:
: TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
:
: TITLE OF INVENTION: GMI-SPECIFIC ALPHA-2 FUCOSYLTRANSFERASE AND USES
:
: TITLE OF INVENTION: THEREOF
:
: FILE REFERENCE: 8511-029
:
: CURRENT APPLICATION NUMBER: US/09/999,672
:
: CURRENT FILING DATE: 2001-10-31
:
: PRIOR APPLICATION NUMBER: US/09/298,886
:
: PRIOR FILING DATE: 1999-04-26
:
: NUMBER OF SEQ ID NOS: 29
:
: SOFTWARE: PatentIn Ver. 2.0
:
: SEQ ID NO 8
:
: LENGTH: 380
:
: TYPE: PRT
:
: ORGANISM: Rattus norvegicus
:
: US-09-999-672-8

```

```
Query Match      100.0%; Score 1921; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.7e-190;
Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

1	STIITHLQORIVKLOP	SEKELPMTQ	SSNTSP	EMRRDSE	QHGNGEL	RGMFTINS	GR 60
2							
3	STIITHLQORIVKLOP	SEKELPMTQ	SSNTSP	EMRRDSE	QHGNGEL	RGMFTINS	GR 82
4							
5							
6							
7	61	LGNOMGEYATL	FALARNGRL	ATIPAS	MHNAL	APIERIS	LPVLSHTAKKI
8							PQNVNHLND 120
9							
10	83	LGNOMGEYATL	FALARNGRL	ATIPAS	MHNAL	APIERIS	LPVLSHTAKKI
11							PQNVNHLND 142
12							
13	121	WMERYRHI	PGHFVRRFTG	YPCSWTFY	HHLRPEIL	KFTLHDH	VREQAQAF
14							RLGLRVNSQ 180
15							
16	143	WMERYRHI	PGHFVRRFTG	YPCSWTFY	HHLRPEIL	KFTLHDH	VREQAQAF
17							RLGLRVNSQ 202
18							
19	181	PSTFVGHVRRG	DVYHMPN	VNWKGV	ADRGYLE	KALDM	FRARYSSP
20							VVVTNSGMACWRE 240
21							
22	203	PSTFVGHVRRG	DVYHMPN	VNWKGV	ADRGYLE	KALDM	FRARYSSP
23							VVVTNSGMACWRE 262
24							

predicted. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	§		DB	ID	Description
			Length				
1	1921	100.0	380	10	US-09-999-672-8	Sequence 8, Appli	
2	1921	100.0	380	12	US-10-040-863-8	Sequence 10, Appli	
3	1896	98.7	353	10	US-09-999-672-10	Sequence 10, Appli	
4	1896	98.7	353	12	US-10-040-863-10	Sequence 10, Appli	
5	1451	75.5	340	10	US-09-051-034A-2	Sequence 2, Appli	
6	1397	72.7	344	10	US-09-999-672-11	Sequence 11, Appli	
7	1397	72.7	344	12	US-10-040-863-11	Sequence 11, Appli	
8	1051.5	54.7	365	9	US-10-105-863-10	Sequence 10, Appli	
9	1051.5	54.7	365	10	US-09-863-475A-6	Sequence 6, Appli	
10	1046.5	54.5	365	10	US-09-051-034A-4	Sequence 4, Appli	
11	1044.5	54.4	365	10	US-09-844-268-13	Sequence 13, Appli	
12	1044.5	54.4	365	10	US-09-844-705-13	Sequence 13, Appli	
13	141	7.3	300	10	US-09-848-838-2	Sequence 2, Appli	
14	99	5.2	496	9	US-10-234-524-25	Sequence 25, Appli	
15	97.5	5.1	431	9	US-09-479-614-14	Sequence 14, Appli	
16	97.5	5.1	496	9	US-09-479-614-2	Sequence 2, Appli	
17	97.5	5.1	496	9	US-09-479-614-29	Sequence 29, Appli	
18	89	4.6	1031	9	US-09-764-868-647	Sequence 647, Appli	
19	89	4.6	2057	9	US-09-815-379-10	Sequence 10, Appli	

QY 241 NINASRGDVVFAGNGIEGSPAKDFALLTQCNTTMTIGTFGWAAYLAGGDTIYLANYT 300  
 Db 263 NINASRGDVVFAGNGIEGSPAKDFALLTQCNTTMTIGTFGWAAYLAGGDTIYLANYT 322  
 QY 301 PDSPLKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCTVAGRAF 358  
 Db 323 PDSPLKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCTVAGRAF 380

## RESULT 2

US-10-040-863-8  
 ; Sequence 8, Application US/10040863  
 ; Patent No. US20020137165A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Eric H. Holmes et al.  
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
 ; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
 ; FILE OF INVENTION: THEREOF  
 ; CURRENT APPLICATION NUMBER: US/10/040,863  
 ; CURRENT FILING DATE: 2001-11-01  
 ; PRIOR APPLICATION NUMBER: 09/298,886  
 ; PRIOR FILING DATE: 1999-04-23  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 8  
 ; LENGTH: 380  
 ; TYPE: PRT  
 ; ORGANISM: Rattus norvegicus  
 ; US-10-040-863-8

Query Match 100.0%; Score 1921; DB 12; Length 380;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-190;  
 Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIIHQRIQVVKLPSEKELPMTQMSSGNTPEMRDSEHQNGELRGMTINSIGR 60  
 Db 23 STIIHQRIQVVKLPSEKELPMTQMSSGNTPEMRDSEHQNGELRGMTINSIGR 82  
 QY 61 LQONQGEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLSHDTAKKIPWQNYHLNDMEER 120  
 Db 83 LQONQGEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLSHDTAKKIPWQNYHLNDMEER 142  
 QY 121 WMEERYRHIPGHVFRTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQ 180  
 Db 143 WMEERYRHIPGHVFRTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQ 202  
 QY 181 PSTFVGVHVRGDYVHVMPNVKGVADRGYLEKALDMFRARYSSPVFVVTSGMWCWRE 240  
 Db 203 PSTFVGVHVRGDYVHVMPNVKGVADRGYLEKALDMFRARYSSPVFVVTSGMWCWRE 262  
 QY 241 NINASRGDVVFAGNGIEGSPAKDFALLTQCNTTMTIGTFGWAAYLAGGDTIYLANYT 300  
 Db 263 NINASRGDVVFAGNGIEGSPAKDFALLTQCNTTMTIGTFGWAAYLAGGDTIYLANYT 322  
 QY 301 PDSPLKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCTVAGRAF 358  
 Db 323 PDSPLKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCTVAGRAF 380

## RESULT 3

US-09-999-672-10  
 ; Sequence 10, Application US/09999672  
 ; Patent No. US20020127655A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Eric H. Holmes et al.  
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
 ; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
 ; FILE OF INVENTION: THEREOF  
 ; FILE REFERENCE: 8511-029  
 ; CURRENT APPLICATION NUMBER: US/09/999,672  
 ; CURRENT FILING DATE: 2001-10-31  
 ; PRIOR APPLICATION NUMBER: US/09/298,886

; PRIOR FILING DATE: 1999-04-26  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: PatentIn ver. 2.0  
 ; SEQ ID NO 10  
 ; LENGTH: 353  
 ; TYPE: PRT  
 ; ORGANISM: Rattus norvegicus  
 ; US-09-999-672-10

Query Match 98.7%; Score 1896; DB 10; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-188;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LQORIVKLOPLSEKELPMTQMSSGNTPEMRDSEHQNGELRGMTINSIGRLGNOM 65  
 Db 1 LQORIVKLOPLSEKELPMTQMSSGNTPEMRDSEHQNGELRGMTINSIGRLGNOM 60  
 QY 66 GEYATLAFALARMNGRLAFIPASMHNALAPIFRISLPVLSHDTAKKIPWQNYHLNDMEER 125  
 Db 61 GEYATLAFALARMNGRLAFIPASMHNALAPIFRISLPVLSHDTAKKIPWQNYHLNDMEER 120  
 QY 126 YRHIPGHVFRTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSTFV 185  
 Db 121 YRHIPGHVFRTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSTFV 180  
 QY 186 GYHVRGDYVHVMPNVKGVADRGYLEKALDMFRARYSSPVFVVTSGMWCWREINAS 245  
 Db 181 GYHVRGDYVHVMPNVKGVADRGYLEKALDMFRARYSSPVFVVTSGMWCWREINAS 240  
 QY 246 RGVVFAGNGIEGSPAKDFALLTQCNTTMTIGTFGWAAYLAGGDTIYLANYTLPDSPF 305  
 Db 241 RGVVFAGNGIEGSPAKDFALLTQCNTTMTIGTFGWAAYLAGGDTIYLANYTLPDSPF 300  
 QY 306 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCTVAGRAF 358  
 Db 301 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCTVAGRAF 353

## RESULT 4

US-10-040-863-10  
 ; Sequence 10, Application US/10040863  
 ; Patent No. US20020137165A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Eric H. Holmes et al.  
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
 ; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
 ; FILE OF INVENTION: THEREOF  
 ; FILE REFERENCE: 8511-029  
 ; CURRENT APPLICATION NUMBER: US/10/040,863  
 ; CURRENT FILING DATE: 2001-11-01  
 ; PRIOR APPLICATION NUMBER: 09/298,886  
 ; PRIOR FILING DATE: 1999-04-23  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 10  
 ; LENGTH: 353  
 ; TYPE: PRT  
 ; ORGANISM: Rattus norvegicus  
 ; US-10-040-863-10

Query Match 98.7%; Score 1896; DB 12; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-188;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LQORIVKLOPLSEKELPMTQMSSGNTPEMRDSEHQNGELRGMTINSIGRLGNOM 65  
 Db 1 LQORIVKLOPLSEKELPMTQMSSGNTPEMRDSEHQNGELRGMTINSIGRLGNOM 60  
 QY 66 GEYATLAFALARMNGRLAFIPASMHNALAPIFRISLPVLSHDTAKKIPWQNYHLNDMEER 125  
 Db 61 GEYATLAFALARMNGRLAFIPASMHNALAPIFRISLPVLSHDTAKKIPWQNYHLNDMEER 120  
 QY 126 YRHIPGHVFRTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSTFV 185

Db 121 YRHLPGHVFTGYPCSWTFYHLLRPKLEFTLHDHVREAAQAFGLRVLNGSQPSTFV 180  
QY 186 GVHVRGDIYVHVPNWKGVDYRGYLEKALDMFRARYSPFVVTNSGMACWRENTNAS 245  
Db 181 GVHVRGDIYVHVPNWKGVDYRGYLEKALDMFRARYSPFVVTNSGMACWRENTNAS 240  
QY 246 RGVVVFAGNGIEGSPAKDFALLTQCNHTIMTIGTGGWAAYLAGGDIYLYANVLPSPF 305  
Db 241 RGVVVFAGNGIEGSPAKDFALLTQCNHTIMTIGTGGWAAYLAGGDIYLYANVLPSPF 300  
QY 306 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHPHLKAKGKAGVTCVYAGRAF 358  
Db 301 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHPHLKAKGKAGVTCVYAGRAF 353

## RESULT 5

US-09-051-034A-2  
; Sequence 2, Application US/09051034A  
; Patent No. US2001005584A1  
; GENERAL INFORMATION:  
; APPLICANT: MCKENZIE, IAN FAROUHAR CAMPBELL  
; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC  
; FILE REFERENCE: 30562.6USWO  
; CURRENT APPLICATION NUMBER: US/09/051,034A  
; CURRENT FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: PCY/AU97/00492  
; PRIOR FILING DATE: 1997-08-01  
; PRIOR APPLICATION NUMBER: 60/024,279  
; PRIOR FILING DATE: 1996-08-21  
; PRIOR APPLICATION NUMBER: P04402  
; PRIOR FILING DATE: 1996-08-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 340  
; TYPE: PRT  
; ORGANISM: Sus Domesticus  
US-09-051-034A-2

Query Match 75.5%; Score 1451; DB 10; Length 340;  
Best Local Similarity 81.9%; Pred. No. 6.9e-142;  
Matches 271; Conservative 26; Mismatches 20; Indels 14; Gaps 3;

QY 1 STIHLQORIVKLOPLSEKELPMYTMQSSNGTSPENRSEQHNGELRCGMFTINSIGR 60  
Db 23 STIHLQORIVKLOPLSEKELPMYTMQSSNGTSPENRSEQHNGELRCGMFTINSIGR 68  
QY 61 LGNOMGEYATLFAALARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPWONYHLND 120  
Db 69 LGNOMGEYATLFAALARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPWONYHLND 128  
QY 121 WMEERYRHIPGHEVFRFTGYPCSWTFYHLLRPKLEFTLHDHVREAAQAFGLRVLNGSQ 180  
Db 129 WMEERYRHIPGHEVFRFTGYPCSWTFYHLLRPKLEFTLHDHVREAAQAFGLRVLNGSQ 188  
QY 181 PSTVGVHVRGDIYVHVPNWKGVDYRGYLEKALDMFRARYSPFVVTNSGMACRE 240  
Db 189 PSTVGVHVRGDIYVHVPNWKGVDYRGYLEKALDMFRARYSPFVVTNSGMACRE 248  
QY 241 NINASRGDVVFAGNGIEGSPAKDFALLTQCNHTIMTIGTGGWAAYLAGGDIYLYANVTL 300  
Db 249 NINASRGDVVFAGNGIEGSPAKDFALLTQCNHTIMTIGTGGWAAYLAGGDIYLYANVTL 308  
QY 301 PDSFPLKVFKEAAFLPEWVGIPADLSPLLK 331  
Db 309 PDSFPLKVFKEAAFLPEWVGIPADLSPLLK 339

## RESULT 6

US-09-999-672-11

; Sequence 11, Application US/09999672  
; Patent No. US20020127655A1  
; GENERAL INFORMATION:  
; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; FILE REFERENCE: 8511-029  
; CURRENT APPLICATION NUMBER: US/09/999,672  
; CURRENT FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: US/09/298,886  
; PRIOR FILING DATE: 1999-04-26  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 344  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-999-672-11

Query Match 72.7%; Score 1397; DB 10; Length 344;  
Best Local Similarity 78.9%; Pred. No. 2.8e-136;  
Matches 262; Conservative 26; Mismatches 32; Indels 12; Gaps 3;

QY 1 STIHLQORIVKLOPLSEKELPMYTMQSSNGTSPENRSEQHNGELRCGMFTINSIGR 60  
Db 23 STIHLQORIVKLOPLSEKELPMYTMQSSNGTSPENRSEQHNGELRCGMFTINSIGR 71  
QY 61 LGNOMGEYATLFAALARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPWONYHLND 120  
Db 72 LGNOMGEYATLFAALARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPWONYHLND 131  
QY 121 WMEERYRHIPGHEVFRFTGYPCSWTFYHLLRPKLEFTLHDHVREAAQAFGLRVLNGSQ 179  
Db 132 WMEERYRHIPGHEVFRFTGYPCSWTFYHLLRPKLEFTLHDHVREAAQAFGLRVLNGSQ 191  
QY 180 QPSTFVGVHVRGDIYVHVPNWKGVDYRGYLEKALDMFRARYSPFVVTNSGMACR 239  
Db 192 QPSTFVGVHVRGDIYVHVPNWKGVDYRGYLEKALDMFRARYSPFVVTNSGMACR 251  
QY 240 ENINASRGDVVFAGNGIEGSPAKDFALLTQCNHTIMTIGTGGWAAYLAGGDIYLYANVTL 299  
Db 252 ENIDASRGDVVFAGNGIEGSPAKDFALLTQCNHTIMTIGTGGWAAYLAGGDIYLYANVTL 311  
QY 300 LPDSFPLKVFKEAAFLPEWVGIPADLSPLLK 331  
Db 312 LPDSFPLKVFKEAAFLPEWVGIPADLSPLLK 343

## RESULT 7

US-10-040-863-11  
; Sequence 11, Application US/10040863  
; Patent No. US20020137165A1  
; GENERAL INFORMATION:  
; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; FILE REFERENCE: 8511-029  
; CURRENT APPLICATION NUMBER: US/10/040,863  
; CURRENT FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: 09/298,886  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 344  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-040-863-11

Query Match 72.7%; Score 1397; DB 12; Length 344;  
Best Local Similarity 78.9%; Pred. No. 2.8e-136;

```
Matches 262; Conservative 26; Mismatches 32; Indels 12; Gaps 3;
QY 1 STIIHQRIYKQLPSKELPMTQMSGNTSPENRRDSQHQNGELRGMTTNSIGR 60
DB 23 STIFHVQORLAKIQAM--WELPV-----QIFVLASTSKALGPSQLRGMTTNAIGR 71
QY 61 LGNQGEVATLALARMNGREAFIPASMHNALAPIRISLPVLHSDTAKKIPWONYHLND 120
DB 72 LGNQGEVATLALAKMNGRPATIPQAQHSHTAPIRITPLVHLNATASRIIPWONYHLND 131
QY 121 WMEERYRHI-PGHVRFYTGPCSWTFYHHLRPEILKEFTLHDVREEAQAFRLGRVNGS 179
DB 132 WMEERYRHIPPGEYVRFYTGPCSWTFYHHLRPEILKEFTLHDVREEAQAFRLGRVNGS 191
QY 180 QSTFTGVHVRGDIYVHPNWKGVADRGYLEKALDMFRARYSPFVVTNSGMWACR 239
DB 192 RPTFTGVHVRGDIYVHPNWKGVADRRYQLQALDMFRARYSLFVVTNSGMWACR 251
QY 240 ENINASRGDWFAGNIGESPAKDFALLTQCNTMTICTGFIWAAYLAGGDTIYLYANT 299
DB 252 ENIDTSHGDVAVGDIYESPAKDFALLTQCNTMTICTGFIWAAYLAGGDTIYLYANT 311
QY 300 LPDSPLKVKPEAEFLPWGIPADLSPLK 331
DB 312 LPDSPLKIKFKPEAEFLPWGTGIAADLSPLK 343
RESULT 8
US-10-105-963-10
; Sequence 10, Application US/10105963
; Publication No. US2003068818A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Denning, Chris
; APPLICANT: Clark, A. John
; TITLE OF INVENTION: Animal Tissue with Carbohydrate Antigens Compatible for Human
; TITLE OF INVENTION: Transplantation and a Carbohydrate Determinant Selection System
; TITLE OF INVENTION: Recombination
; FILE REFERENCE: 731/002
; CURRENT APPLICATION NUMBER: US/10/105,963
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: US 60/277,811
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-105-963-10
Query Match 54.7%; Score 1051.5; DB 9; Length 365;
Best Local Similarity 63.3%; Pred. No. 1.9e-100;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 30 GNTPEMRDRDSEHQNGELRGMTTNSIGRLGNQGEVATLALARMNGRLAFIPASMH 89
DB 59 GTAMGNASSCCPQH-PASLSGTWTVVPGNRCFNQMGQVATLLAQLNGRRAFILPAMH 117
QY 90 NALAPIRISLPVLHSDTAKKIPWONYHLNDWMEERYRHIPGHVRFYTGPCSWTFYHHL 149
DB 118 AALAPVFRITPLVLAPEVDSRIIPWRELQLHDWMSSEYADLRDPLKLSGPGCSWTFYHHL 177
QY 150 RPEILKEFTLHDVREEAQAFRLGRV--NGSOPSTFVGVHVRGDIYVHPNWKGVVA 207
DB 178 REQIRREFTLHDVREEAQAFRLGRV--NGSOPSTFVGVHVRGDIYVHPNWKGVVA 237
QY 208 DRGYLEKALDMFRARYSPFVVTNSGMWACRENINASRGDVFVAGNIGESPAKDFALL 267
DB 238 DSAYLRQMDWFRARHAPVTVVNSGMWCKENIDTSCQGDVTFAGDQGTATWKFQFALL 297
QY 268 TCNHTMTICTGFIWAAYLAGGDTIYLYANTPLDPSPLKVKPEAEFLPWGTGIPADLS 327
```

```
DB 298 TCNHTMTICTGFIWAAYLAGGDTIYLYANTPLDPSPLKVKPEAEFLPWGTGIPADLS 357
QY 328 PLKALTP 335
DB 358 PLWILAKP 365
RESULT 9
US-09-863-475A-6
; Sequence 6, Application US/09863475A
; Patent No. US20020102688A1
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/09/863,475A
; APPLICATION NUMBER: 07/914,281
; FILING DATE: 20-JUL-1992
; PRIORITY INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT GR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-863-475A-6
Query Match 54.7%; Score 1051.5; DB 10; Length 365;
Best Local Similarity 63.3%; Pred. No. 1.9e-100;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 30 GNTPEMRDRDSEHQNGELRGMTTNSIGRLGNQGEVATLALARMNGRLAFIPASMH 89
DB 59 GTAMGNASSCCPQH-PASLSGTWTVVPGNRCFNQMGQVATLLAQLNGRRAFILPAMH 117
QY 90 NALAPIRISLPVLHSDTAKKIPWONYHLNDWMEERYRHIPGHVRFYTGPCSWTFYHHL 149
DB 118 AALAPVFRITPLVLAPEVDSRIIPWRELQLHDWMSSEYADLRDPLKLSGPGCSWTFYHHL 177
QY 150 RPEILKEFTLHDVREEAQAFRLGRV--NGSOPSTFVGVHVRGDIYVHPNWKGVVA 207
DB 178 REQIRREFTLHDVREEAQAFRLGRV--NGSOPSTFVGVHVRGDIYVHPNWKGVVA 237
QY 208 DRGYLEKALDMFRARYSPFVVTNSGMWACRENINASRGDVFVAGNIGESPAKDFALL 267
```



Db 238 DSAYLRQAMDFAHRAEPVFTVTSNGMEWKENIDTSQGDVTFAGDGOEATPKDFALL 297  
QY 268 TQCNFTTMTGTFGWAAYLAGDGTIYANITLPSDPFLKVKFPEAAFLPEWGPADLS 327  
Db 238 TQCNFTTMTGTFGWAAYLAGDGTIYANITLPSDPFLKVKFPEAAFLPEWGINADLS 357  
QY 328 PLLKALTP 335  
Db 358 PLWTLAKP 365

## RESULT 10

US-09-051-034A-4  
; Sequence 4, Application US/09051034A  
; Patent No. US20010055584A1

## GENERAL INFORMATION:

; APPLICANT: MCKENZIE, IAN FARQUHAR CAMPBELL

; APPLICANT: SANDRIN, MAURO SERGIO

; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC

; FILE REFERENCE: GLYCOSYLTRANSFERASE

; CURRENT APPLICATION NUMBER: US/09/051.034A

; PRIOR FILING DATE: 1998-03-31

; PRIOR APPLICATION NUMBER: PCT/AU97/00492

; PRIOR FILING DATE: 1997-08-01

; PRIOR APPLICATION NUMBER: 60/024,279

; PRIOR FILING DATE: 1996-08-21

; PRIOR APPLICATION NUMBER: P01402

; PRIOR FILING DATE: 1996-08-02

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 4

; LENGTH: 365

; TYPE: PRT

; ORGANISM: Sus Domesticus

; US-09-051-034A-4

Query Match 54.5%; Score 1046.5; DB 10; Length 365;  
Best Local Similarity 62.6%; Pred. No. 6.2e-100;  
Matches 194; Conservative 38; Mismatches 75; Indels 3; Gaps 2;

QY 22 PMTQMSSGNTSPEMRRDSEQHNGELRGMTINSIGRLGNQMGYATLAFALARNGR 81  
Db 51 PVAIFCLAGTPVHPNASDSCPKH-PASFSGTWTIYPDGREGNQMGYATLLALAQLNGRQ 109  
QY 82 AFIPASMHNALAPIRISLPVLHSDTAKKIPQNYHLNDWMEERYRHIPGHVFTGYPC 141  
Db 110 AFIQPAMHVALPVPFRITLPVLAPEVDHRAHPWELELDHDMSEDAHLKEPWLKLTGFC 169  
QY 142 SWTFYHHLRPEILKEFTLHDHVREREAQAFRLGRV--NGSQPSTFVGHVRRGDYVHYMP 199  
Db 170 SWTFYHHLRPEILKEFTLHDHVREREAQAFRLGRV--NGSQPSTFVGHVRRGDYVHYMP 229  
QY 200 NWKGVADRGVLEKALDMFRARYSPVFTVTSNGMARENINASRGDVFAGNGIEGS 259  
Db 230 KWKGVGVDGALQOAMDWRARYEAPVFTVTSNGMEWKCKNIDTSRGDVFAGDGEAA 289  
QY 260 PAKDFALLTQCNTTMTGTGFWAAYLAGDGTIYANITLPSDPFLKVKFPEAAFLPEW 319  
Db 290 PAKDFALLTQCNTTMTGTGFWAAYLAGDGTIYANITLPSDPFLKVKFPEAAFLPEW 349  
QY 320 VGIPADLSPL 329  
Db 350 VGINADLSPL 359

## RESULT 11

US-09-844-268-13

; Sequence 13, Application US/09844268

; Patent No. US20020129395A1

## GENERAL INFORMATION:

; APPLICANT: BOSWORTH, BRAD

; APPLICANT: VOGELI, PETER  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY  
; FILE REFERENCE: RESISTANT TO F18 E. COLI ASSOCIATED DISEASES  
; CURRENT APPLICATION NUMBER: US/09/844,268  
; CURRENT FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: 09/443,766  
; PRIOR FILING DATE: 1999-11-19  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 365  
; TYPE: PRT  
; ORGANISM: Porcine  
; US-09-844-268-13

Query Match 54.4%; Score 1044.5; DB 10; Length 365;  
Best Local Similarity 62.6%; Pred. No. 1e-99;  
Matches 194; Conservative 38; Mismatches 75; Indels 3; Gaps 2;

QY 22 PMTQMSSGNTSPEMRRDSEQHNGELRGMTINSIGRLGNQMGYATLAFALARNGR 81  
Db 51 PVAIFCLAGTPVHPNASDSCPKH-PASFSGTWTIYPDGREGNQMGYATLLALAQLNGRQ 109  
QY 82 AFIPASMHNALAPIRISLPVLHSDTAKKIPQNYHLNDWMEERYRHIPGHVFTGYPC 141  
Db 110 AFIQPAMHVALPVPFRITLPVLAPEVDHRAHPWELELDHDMSEDAHLKEPWLKLTGFC 169  
QY 142 SWTFYHHLRPEILKEFTLHDHVREREAQAFRLGRV--NGSQPSTFVGHVRRGDYVHYMP 199  
Db 170 SWTFYHHLRPEILKEFTLHDHVREREAQAFRLGRV--NGSQPSTFVGHVRRGDYVHYMP 229  
QY 200 NWKGVADRGVLEKALDMFRARYSPVFTVTSNGMARENINASRGDVFAGNGIEGS 259  
Db 230 KWKGVGVDGALQOAMDWRARYEAPVFTVTSNGMEWKCKNIDTSRGDVFAGDGEAA 289  
QY 260 PAKDFALLTQCNTTMTGTGFWAAYLAGDGTIYANITLPSDPFLKVKFPEAAFLPEW 319  
Db 290 PAKDFALLTQCNTTMTGTGFWAAYLAGDGTIYANITLPSDPFLKVKFPEAAFLPEW 349  
QY 320 VGIPADLSPL 329  
Db 350 VGINADLSPL 359

## RESULT 12

US-09-844-705-13

; Sequence 13, Application US/09844705

; Patent No. US2002013836A1

## GENERAL INFORMATION:

; APPLICANT: BOSWORTH, BRAD

; APPLICANT: VOGELI, PETER

; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY

; FILE REFERENCE: RESISTANT TO F18 E. COLI ASSOCIATED DISEASES

; CURRENT APPLICATION NUMBER: US/09/844,705

; CURRENT FILING DATE: 2001-04-27

; PRIOR APPLICATION NUMBER: 09/443,766

; PRIOR FILING DATE: 1999-11-19

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 13

; LENGTH: 365

; TYPE: PRT

; ORGANISM: Porcine

; US-09-844-705-13

Query Match 54.4%; Score 1044.5; DB 10; Length 365;  
Best Local Similarity 62.6%; Pred. No. 1e-99;  
Matches 194; Conservative 38; Mismatches 75; Indels 3; Gaps 2;

QY 22 PMTQMSSGNTSPEMRRDSEQHNGELRGMTINSIGRLGNQMGYATLAFALARNGR 81  
Db 51 PVAIFCLAGTPVHPNASDSCPKH-PASFSGTWTIYPDGREGNQMGYATLLALAQLNGRQ 109  
QY 82 AFIPASMHNALAPIRISLPVLHSDTAKKIPQNYHLNDWMEERYRHIPGHVFTGYPC 141  
Db 110 AFIQPAMHVALPVPFRITLPVLAPEVDHRAHPWELELDHDMSEDAHLKEPWLKLTGFC 169  
QY 142 SWTFYHHLRPEILKEFTLHDHVREREAQAFRLGRV--NGSQPSTFVGHVRRGDYVHYMP 199  
Db 170 SWTFYHHLRPEILKEFTLHDHVREREAQAFRLGRV--NGSQPSTFVGHVRRGDYVHYMP 229  
QY 200 NWKGVADRGVLEKALDMFRARYSPVFTVTSNGMARENINASRGDVFAGNGIEGS 259  
Db 230 KWKGVGVDGALQOAMDWRARYEAPVFTVTSNGMEWKCKNIDTSRGDVFAGDGEAA 289  
QY 260 PAKDFALLTQCNTTMTGTGFWAAYLAGDGTIYANITLPSDPFLKVKFPEAAFLPEW 319  
Db 290 PAKDFALLTQCNTTMTGTGFWAAYLAGDGTIYANITLPSDPFLKVKFPEAAFLPEW 349  
QY 320 VGIPADLSPL 329  
Db 350 VGINADLSPL 359

Db 51 PVATFCLAGTVPHNDSADSCPKH--PASFSQGTWTYPDGRFGNQMGQYATLLAQLNSRQ 109  
QY 82 APTPASHNALAPIFRISLVLHSDTAKKIPWQNYHLNDWMEERYRHPGHFVFTCYPC 141  
Db 110 AFTQPAHVALVAFVFTLVLAPVDRHAPWSELELDHDMSEDYAHLKFEWKLGTGEP 169  
QY 142 SWTFYHHLRPEILKEFTLHDHVREREAQAFRLGRV--NGSQPSRFVGVHVRGGRDYVHVMP 199  
Db 170 SWTFYHHLRPEILKEFTLHDHVREREAQAFRLGRV--NGSQPSRFVGVHVRGGRDYVHVMP 229  
QY 200 NVKGVVADRGYLEKALDMFRARYSSPVFVVTNSGMWCRENINASRGDVVYFAGNGIEGS 259  
Db 230 KRWKGVGDGRYLLQQAQMDWFRARYEAPVVFVVTNSGMWCRENINASRGDVVYFAGNGIEGS 289  
QY 260 PAKDFALLTCQNHNTIMTIGFIAWAALAGDGTIYLANITLPSDFLKVPKPEAAFLPEW 319  
Db 290 PARDFALLVCQNHNTIMTIGFIAWAALAGDGTIYLANITLPSDFLKVPKPEAAFLPEW 349  
QY 320 VGIPADLSPL 329  
Db 350 VGINADLSPL 359  
RESULT 13  
US-09-848-838-2  
; Sequence 2, Application US/09848838  
; Patent No. US20020037570A1  
; GENERAL INFORMATION:  
; APPLICANT: Taylor, biane  
; APPLICANT: Wang, Ge  
; APPLICANT: Palcio, Monica  
; TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE  
; FILE REFERENCE: 07254-061002  
; CURRENT APPLICATION NUMBER: US/09/848,838  
; CURRENT FILING DATE: 2001-05-03  
; PRIOR APPLICATION NUMBER: US 09/433,598  
; PRIOR FILING DATE: 1999-11-02  
; PRIOR APPLICATION NUMBER: US 60/107,268  
; PRIOR FILING DATE: 1998-11-04  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 300  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori  
US-09-848-838-2

Query Match 7.3%; Score 141; DB 10; Length 300;  
Best Local Similarity 24.0%; Pred. No. 3e-06;  
Matches 81; Conservative 33; Mismatches 125; Indels 98; Gaps 14;  
QY 59 GRLGNQMGEXA-----TLFALARMNGRLAFIPASMHNALAPIFRIS-- 99  
Db 10 GGLGNQMGQFAFAKSLQKHLNTPVLLDTTSDNSNRKQLELPIDIPYANAKETIAKM 69  
QY 100 --LPVLHSDTAKKIPWQNYHLNDWMEERYRHI-PGHVFRFTGYPCSWTFYHHLRPEILKE 156  
Db 70 QHLFLKLYRDALKYIGEDRVSOEIVFEYEPKLLKPSRLTYFFGYFQDPDYDAISLLIKOT 129  
QY 157 FTL-----HDHVREREAQAFRLGRVNSQPSSTFVGVHVRGGRDYVHVMPNVKGVVA 207  
Db 130 FTLPPLPPPPENKNNKKEEYQKLS--LILAANKSVF--VHIRRGDYVGI-----GCQL 179  
QY 208 DRGYLEKALDMFRARYSSPVFVVTNSGMWCRENINASRGDVVYFAGNGIEGSPAKDFA-- 265  
Db 180 GIDYQKALFEMAKRVNMEFLV-----FCE-----DLKFTQMLDLGYPTDMTR 225  
QY 266 -----LLTQCQNHNTIMTIGFIAWAALAGDGTIYLANITLPSDFLKVPK 310  
Db 226 DKBEAYWDMLLMQSKGHGIANSTYSWAAYL-----MEN---PEKIIIGPKHHLFG 275  
QY 311 PEAFLPEWGVIGFADLSPLLKALTTPACPRSHHFLKAK 347

Db 276 HENLCKRWVKI-----ESHFEVKSQ 296  
RESULT 14  
US-10-214-524-25  
; Sequence 25, Application US/10214524  
; Publication No. US20030073142A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Swei-Shen Alex  
; APPLICANT: Yang, Yong-Min  
; APPLICANT: Barankiewicz, Theresa J.  
; APPLICANT: Chen, Zhong  
; TITLE OF INVENTION: IMMUNOGLOBULIN E VACCINES AND METHODS OF USE THEREOF  
; FILE REFERENCE: ICE-00101-P.1.1  
; CURRENT APPLICATION NUMBER: US/10/214,524  
; CURRENT FILING DATE: 2002-08-08  
; PRIOR APPLICATION NUMBER: 60/312,120  
; PRIOR FILING DATE: 2001-08-13  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 25  
; LENGTH: 496  
; TYPE: PRT  
; ORGANISM: Cat (Felis catus)  
US-10-214-524-25  
Query Match 5.2%; Score 99; DB 9; Length 496;  
Best Local Similarity 25.0%; Pred. No. 0.14;  
Matches 55; Conservative 21; Mismatches 84; Indels 60; Gaps 12;  
QY 131 GHFVRFTGYPCSWTFYHHLRPEILKEFTLHDHVREREAQAFRLGRVNSQPSSTFVGVHVR 190  
Db 249 GEWVSQKTYTCQYI-----QGETFEDHARKCHESPRGVSTYLSPPSL----- 293  
QY 191 RGVYHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTNSGM--AWCRENINASRGD 248  
Db 294 -DLVYHKSPIK-TCLVVD-----LANTDGMILTWSREN----- 324  
QY 249 VVFAGNGIEGSPAKDFALLTCQNHNTIMTIGFIAWAALAGDGTIYLANITLPSD---P 305  
Db 325 ---GESVHPDP---VWKTKYNGTITVTSTLPVDATDWWVEGET-YOCKVTHPDLPKDIV 376  
QY 306 LKVEK-PEAAFLPE-WVGIPADLSPLK-ALTPACPRSHF 342  
Db 377 RSIKAPGRFPPEVYVFLPPEGEPTKDKVTLTCLQNF 416  
RESULT 15  
US-09-479-614-14  
; Sequence 14, Application US/09479614  
; Publication No. US20030013183A1  
; GENERAL INFORMATION:  
; APPLICANT: McCall, Catherine  
; APPLICANT: Weber, Eric  
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods  
; FILE REFERENCE: P-1047  
; CURRENT APPLICATION NUMBER: US/09/479,614  
; CURRENT FILING DATE: 2000-01-07  
; EARLIER APPLICATION NUMBER: 60/115,033  
; EARLIER FILING DATE: 1999-01-07  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Felis catus  
US-09-479-614-14  
Query Match 5.1%; Score 97.5; DB 9; Length 431;  
Best Local Similarity 25.0%; Pred. No. 0.16;  
Matches 52; Conservative 19; Mismatches 78; Indels 59; Gaps 11;  
QY 131 GHFVRFTGYPCSWTFYHHLRPEILKEFTLHDHVREREAQAFRLGRVNSQPSSTFVGVHVR 190

```

      | : | | | | : | | | | : | | : | | : | |
Db 184 GEVVSOKTYTCQVY-----QGTFEDHARKCTESDPRGVSTYLSPPSL----- 228
QY 191 RGDYVHVMPNVKGVYADRGYLEKALDMFARYSSPVFVTSNGM--AWCRENINASRGD 248
      | | | | : | | : | | : | | : | | : | |
Db 229 -DLYVHKSPKI--TCLVVD-----LANTDGMILTWSREN----- 259
QY 249 VVFAGNGIEGSPAKDFALLTQC�HTIMTIGTGIWAAYLAGGDTIYLANVTLPDSP---F 305
      | : | | | | | | : | | | | : | | | | |
Db 260 ----GESVHPDP---MVKKIQYNGTIVTVSTLFPVDATDWVEGET-YQCKVTHPDLPKDIV 311
QY 306 LKVKF-PEAAFLPE-WVGIPADLSPLK 331
      : | | | | : | | : | |
Db 312 RSIKAPGRFRFPPEVYVFLPPEGEPKIK 339
```

Search completed: May 27, 2003, 15:25:42  
Job time : 16.4676 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:00:45 ; Search time 12.9728 Seconds  
(without alignments)  
2652.940 Million cell updates/sec

Title: US-10-040-863-8\_COPY\_23\_380

Perfect score: 1921

Sequence: 1 STIIHLOQRIVKLOPLSEKE.....RSHFLKAKGVTCYVAGRAF 358

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1407.5	73.3	343	2 A56098	alpha(1,2)fucosylt
2	1296.5	67.5	354	2 B56392	beta-galactoside a
3	1086	56.5	373	2 A56392	beta-galactoside a
4	1051.5	54.7	365	2 A36047	galactoside 2-alph
5	846	44.0	159	2 S46494	galactoside 2-alph
6	581	30.2	142	2 S51582	galactoside 2-alph
7	214.5	11.2	281	2 T44328	hypothetical prote
8	183	9.5	443	2 H87911	protein B0205.4 [i
9	170.5	8.9	363	2 T20745	hypothetical prote
10	154	8.0	348	2 T31916	hypothetical prote
11	153.5	8.0	335	2 T32294	hypothetical prote
12	139.5	7.3	383	2 T20572	hypothetical prote
13	138.5	7.2	500	2 T22068	hypothetical prote
14	133.5	6.9	299	2 H71976	probable alpha(1,2
15	131	6.8	365	2 T23309	hypothetical prote
16	127	6.6	388	2 T32307	hypothetical prote
17	127	6.6	392	2 T23334	hypothetical prote
18	124	6.5	353	2 T25390	hypothetical prote
19	120	6.2	625	2 T21051	hypothetical prote
20	119	6.2	434	2 T26275	hypothetical prote
21	118.5	6.2	381	2 T15140	hypothetical prote
22	117.5	6.1	365	2 T33253	hypothetical prote
23	111	5.8	395	2 T32309	hypothetical prote
24	109.5	5.7	371	2 T32692	hypothetical prote
25	109	5.7	277	2 T25307	hypothetical prote
26	98.5	5.1	317	2 T33887	hypothetical prote
27	96.5	5.0	355	2 T34405	hypothetical prote
28	93.5	4.9	348	2 T02798	hypothetical prote
29	93	4.8	597	2 G87258	acyl-CoA dehydroge

hypothetical prote  
transforming prote  
probable myp-prote  
glycoprotein D pre  
fibrillin precurs  
probable membrane  
peptide synthetase  
hypothetical prote  
myosin X - human  
glycoprotein D pre  
modular adaptor Gr  
acyl carrier prote  
procollagen-lysine  
growth factor rece  
growth factor rece  
NADH2 dehydrogenas

## ALIGNMENTS

### RESULT 1

A56098

alpha(1,2)fucosyltransferase Sec2, long form - human

C:Species: Homo sapiens (man)

C>Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 21-Jul-2000

C:Accession: A56098

R:Kelly, R.J.; Rouquier, S.; Giorgi, D.; Lennon, G.G.; Lowe, J.B.

J. Biol. Chem. 270, 4640-4649, 1995

A:Title: Sequence and expression of a candidate for the human secretor blood group al  
tes with the non-secretor phenotype.

A:Reference number: A56098; MUID:95181460; PMID:7876235

A:Accession: A56098

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-343 <REL>

A:Cross-references: GB:U17894; NID:g687618; PIDN:AAC24453.1; PID:g687619

C:Genetics:

A:Gene: GDB:FUT2; SE

A:Cross-references: GDB:120619; OMIM:182100

A:Map position: 19q13.3-19q13.3

C:Keywords: glycoprotein; Golgi apparatus; transmembrane protein

Query Match 73.3%; Score 1407.5; DB 2; Length 343;  
Best Local Similarity 79.2%; Pred. No. 3.3e-110;  
Matches 262; Conservative 26; Mismatches 32; Indels 11; Gaps 2;

QY 1 STIIHLOQRIVKLOPLSEKELPMTQMSSQNTSEPMRRDSEQHGNGELRGMTINSIGR 60

Db 23 STIFHYQORLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGMWTINAIGR 71

QY 61 LGNQMGAYATLPAARMNGRLAIPASMHNALAPIFRISLPVLHSDTAKKIPQWNYHLND 120

Db 72 LGNQMGAYATLYALAKMNGRPAFIPAQMHSITLAPIFRITLPVLHSAFASIPQWNYHLND 131

QY 121 WMERYRHIPGHEVTRGTGPCSWTFYHHLRPEILKFTLHDHVREAAQAFRLGRVNGSQ 180

Db 132 WMEERYRHIPGHEVTRGTGPCSWTFYHHLRPEILKFTLHDHVREAAQAFRLGRVNGSR 191

QY 181 PSTFVGHVRRGDYVHVMNPNVWKGVVADRGYLEKALDMFRARYSSPVFVVTSSNGMAWCRE 240

Db 192 PGTFVGHVRRGDYVHVMNPNVWKGVVADRGYLEKALDMFRARYSSPVFVVTSSNGMAWCRE 251

QY 241 NINASGDVVVFAGNEGSPAKDFALLTQCNHNTIMTIGFGIWAAYLAGDGTIYLANITL 300

Db 252 NIDTSGDVVVFAGNEGSPAKDFALLTQCNHNTIMTIGFGIWAAYLAGDGTIYLANITL 311

QY 301 PDSFPLKVKPEAAFLPEWVGIPADLSPLLK 331

Db 312 PDSFPLKVKPEAAFLPEWVGIPADLSPLLK 342

### RESULT 2

Db 79 -----GIWTIHPDGRFGNQGYATLLALAO LNGRRAFILPAMHAA LAPVFRITL 128

R;Piau, J.P.; Labarrière, N.; Dabouis, G.; Denis, M.G.  
Picchem T 300 532-533 1994

R;Plau, J.P.; Labarrière, N.; Dabouis, G.; Denis, M.G.  
Picobom T 300 533-535 1994

A;title: Evidence for two distinct alpha-(1,2)-fucosyltransferase genes differentially expressed in the developing mouse brain  
A;Reference number: S46493; MID:94260382; PMID:8010942  
A;Accession: S46494  
A;Molecule type: mRNA  
A;Residues: 1-159 <PIA>  
A;Cross-references: EMBL:L26010; NID:G414816; PID:AAB41515.1; PID:G414817  
C;Genetics:  
A;Gene: FTB  
C;Keywords: glycosyltransferase; hexosyltransferase

[illegible]

RESULT 6  
S51582  
galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 07-May-1995 #sequence\_revision 17-Nov-1995 #text\_change 08-Oct-1999  
C;Accession: S51582; S46493  
R;Plau, J.P.; Labarrière, N.; Dabouis, G.; Denis, M.G.  
submitted to the EMBL Data Library, November 1993  
A;Reference number: S51582  
A;Accession: S51582  
A;Status: preliminary  
A:Molecule type: mRNA  
A;Residues: 1-142 <PIA>  
A;Cross-references: EMBL:L26009; NID:g414814; PID:g554438  
R;Plau, J.P.; Labarrière, N.; Dabouis, G.; Denis, M.G.  
Biochem. J. 300, 623-626, 1994  
A;Title: Evidence for two distinct alpha-(1,2)-fucosyltransferase genes differentially  
expressed in the rat  
A;Reference number: S46493; MUID:94280382; PMID:8010942  
A;Accession: S46493  
A;Status: preliminary  
A:Molecule type: mRNA  
A;Residues: 9-135 <PI2>  
A;Cross-references: EMBL:L26009  
C;Keywords: glycosyltransferase; hexosyltransferase

	Query Match	30.2%;	Score 581;	DB 2;	Length 142;
	Best Local Similarity	73.2%;	Pred. No. 1.8e-41;		
	Matches 104;	Conservative 15;	Mismatches 21;	Indels 2;	Gaps 1;
Qy	140	PCSTFTFYHLLRPETLKFTFLHDHVBREAQAFRLGLRYN--CSQSPSTFGVYHVRGDIYHV	197		
D6	1	PCSTFTFHLREQRRFTFLHDHVBREAQRLSLGLRGIPAGIRPTTYGVYHVRGDIYLV	60		
Qy	198	MPNWKGVGVADRGYLEKALDMFARARYSPFVVTNSGMACWREINASRGDVPVAGNCIE	257		
D6	61	MPNWKGVGVDRYLYQKAMDWFARHKDPIFVVTNSGMWCLNIDTSHGVDVPVAGNCQE	120		
Qy	258	GSPAKDFALLTCOCHTMTICT	279		
D6	121	GTPGKDFALLTCOCHTMTICT	142		

```

RESULT 7
T44328
hypothetical protein wblA [imported] - Vibrio cholerae
C:Species: Vibrio cholerae

```

C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 28-Jul-2000  
C;Accession: T44328  
R;Yamasaki, S.; Shimizu, T.; Hoshino, K.; Ho, S.T.; Shimada, T.; Nair, G.B.; Takeda, Gene 237, 321-332, 1999  
A;Title: The genes responsible for O-antigen synthesis of *Vibrio cholerae* O139 are cl  
A;Reference number: 222749; MUID:99453293; PMID:10521656  
A;Accession: T44328  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-281 <YAM>  
A;Cross-references: EMBL:AB012957; NID:q4115683; PIDN:BA033632.1; PID:g3721682  
A;Experimental source: strain O22  
C;Genetics:  
A;Note: wblA  
C;Superfamily: *Vibrio cholerae* hypothetical protein wblA

[illegible]

RESULT 8  
H87911  
protein B0205.4 [imported] - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C/Accession: H87911  
R/Anonymous, The C. elegans Sequencing Consortium.  
Science 283, 201-2018, 1998  
A/Title: Genome sequence of the nematode *C. elegans*: a platform for investigating bio  
A/Reference number: A75000; MUID: 99069613; PMID: 9851916  
A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C-  
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;  
A/Accession: H87911  
A/Status: Preliminary  
A/Molecule type: DNA  
A/Residues: 1-443 <STO>  
A/Cross-references: GB:chr.1; PIDN:AC16388.1; PID:G3150470; GSPDB:GN00019; CESP:B0205.4  
A/Note: contains weak similarity to tucosyltransferases  
C/Genetics:  
A/Gene: B0205.4  
A/Map position: 1

[illegible]

Db 163 ASKITLITSCAYRNLSLTLFNDRI-----IERIDGYFONFRYFHPDQKIVKKLFTFM 216  
QY 161 DHVREAOAF-----LRGLR 175  
Db 217 DPKRVKSWNIYNIHPTNRHKPEKSTVSGFVTFQLRVDFLENVGLSLTVARNAR 276  
QY 176 V-----NGSQ-----PSTFVGVBHVRGVDVHV--MPNVKGVVADRGYLEKAL 216  
Db 277 VIETNVANDQALEPEEDAFKTMVGVHVRGMDISMNSNRHGHVDVDTIEYKRAI 336  
QY 217 DMFRARYSPFVWVTSNGMAWCREINASRGDVPFAGNGIEGSPAKDFALLTQCNTHTMT 276  
Db 337 QGTSKIYENVAFLICSDNVAVARNKILGKETLHF--CPGPREVDMAILKSCDSVITS 393  
QY 277 IGTFGWAAYL--AGGDTIYLANVTLTDPSPFLKYFKPEAFLPEWVGI 322  
Db 394 TGTGWSAYLNVNASPDVYIYKHPAGSVMEKMTKTEYFLKSWTAL 442

RESULT 9  
T20745  
hypothetical protein F11A5.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000  
C:Accession: T20745  
R:Gardner, A.  
submitted to the EMBL Data Library, March 1997  
A:Reference number: Z19319  
A:Accession: T20745  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-363 <SWIL>  
A:Cross-references: EMBL:292830; PIDN:CAB07352.1; GSPDB:GN00023; CESP:F11A5.5  
A:Experimental source: clone F11A5  
C:Genetics:  
A:Gene: CESP:F11A5.5  
A:Map position: 5  
A:Introns: 42/1; 55/2; 136/3; 175/3; 210/2; 243/1; 299/1; 330/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 8.9%; Score 170.5; DB 2; Length 363;  
Best Local Similarity 23.4%; Pred. No. 1.4e-06;  
Matches 73; Conservative 45; Mismatches 101; Indels 93; Gaps 17;

QY 57 SIGRLGNMGCVATLAFALARMNGLA--FIPASMH-----NALAPIRISLPVLHSDT 107  
Db 78 SSSRLGNHLFELASVLSISRELQVPTFIENCYHEKMHEDSNTLIPGLMNHLLINGSV 137  
QY 108 AKKIPQNVH-----LNDWMEERYRHIPGHVFTGYPCSWTFYHHLRPEILKEF 157  
Db 138 PSSVRVKFQKCCFTDDPSLLDNYEDEYHLTG-----THYQ-SWKYFSHMRNELI--- 188  
QY 158 TLHDHVREEAQAFRLGLRVNGSQPSTFVG-VHVRGDVYHVMNVKGV-VADRGYLEKA 215  
Db 189 ---GILTKTENY---MDLPKGGENTFTCVHVRGDVFLRV-----GPHVADENPIRS 236  
QY 216 LDMFRARYSPFVWVTSNGMAWCREINASRGDVPFAGNGIE----- 257  
Db 237 LNLISQVAKRANAT-----VEFGDDYEFMSDLRNPSTKINAFVQ 278  
QY 258 GSPAKDFALL-TQCNHTMTI--GTFGIWAAYLAGDITIYLANVTLPDSPFLK----- 307  
Db 279 NSPADLLYAKNSCDVLLTAHSTFGWNGVFSKGNRY---YT--DTQFTKDWILETG 333  
QY 308 VKPEAARLPEW 319  
Db 334 EFISEDYILPHW 345

RESULT 10  
T31916  
hypothetical protein C17A2.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jul-2000  
C:Accession: T31916  
R:Sammons, L.; Wohldmann, P.  
submitted to the EMBL Data Library, July 1997  
A:Description: The sequence of C. elegans cosmid C17A2.  
A:Reference number: Z21098  
A:Accession: T31916  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-348 <SAM>  
A:Cross-references: EMBL:AF016654; PIDN:AAB66030.1; GSPDB:GN00020; CESP:C17A2.4  
A:Experimental source: strain Bristol N2; clone C17A2  
C:Genetics:  
A:Gene: CESP:C17A2.4  
A:Map position: 2  
A:Introns: 94/3; 133/3; 168/2; 272/1  
C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 8.0%; Score 154; DB 2; Length 348;  
Best Local Similarity 22.8%; Pred. No. 3.3e-05;  
Matches 63; Conservative 53; Mismatches 102; Indels 58; Gaps 15;

QY 57 STGRGNCMGCVATLAFALARMNGL--LAFIPASMHNALAPIRISLP-----VLHSDT 107  
Db 36 SPSRLGNHIFEFASLFGLSERLHRTPLFIVNEFHOKMLDETRKVMPLGVEKFTVINGSL 95  
QY 108 AKKI---PWQ-----NYHLNDWMEERYRHIPGHVFTGYPCSWTFYHHLRPEILKEF 157  
Db 96 PHSIKOTPFQKVCCHENPEILEKIDEXYHLTGMYQ-----SWKYFPNMRQELL--- 146  
QY 158 TLHDHVREEAQAFRLGLRVNGSQPSTFVG-VHVRGDVYHVMNVKGV-VADRGYLEKAL 216  
Db 147 ---DFLDDSSQDF--GNLPRSNQRHTVTCVHARRGDFDV-----GFOAADPDFIRNSV 195  
QY 217 DMFRARYSPV-----FVYTSNGMAWCR---EN--INASRGDVPFAGNGI--EGSPA 261  
Db 196 KYIAENFIPEIEYKVKHRRKVVIFGDLFEMLSLFNSVSTDEPEYMPFAEYIISONSPA 255  
QY 262 KDFALLTG--CNHTMTI--GTFGIWAAYLAGDITIY 294  
Db 256 EDLTSYKQNCDIVLISAPKSTFGWIGYFSKGNKVF 291

RESULT 11  
T32294  
hypothetical protein K06H6.6 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 20-Jun-2000  
C:Accession: T32294  
R:David, M.; Wohldmann, P.; Bauer, C.; Clarke, K.  
submitted to the EMBL Data Library, September 1997  
A:Description: The sequence of C. elegans cosmid K06H6.  
A:Reference number: Z21147  
A:Accession: T32294  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-335 <DAV>  
A:Cross-references: EMBL:AF024500; PIDN:AAB70365.1; GSPDB:GN00023; CESP:K06H6.6  
A:Experimental source: strain Bristol N2; clone K06H6  
C:Genetics:  
A:Gene: CESP:K06H6.6  
A:Map position: 5  
A:Introns: 17/1; 32/2; 58/1; 111/3; 150/3; 184/2; 218/1; 243/3; 272/1; 303/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 8.0%; Score 153.5; DB 2; Length 335;  
Best Local Similarity 22.8%; Pred. No. 3.5e-05;  
Matches 77; Conservative 50; Mismatches 118; Indels 93; Gaps 19;

QY 29 SGNITESPNERDSEQHNGELRGMTTINSIGRLGNMGCVATLAFALARMNGLA--FIPA 86  
Db 27 SATAEIPPLRQESFLIMKKRLSRMA--PTARLGNHMFELAAVLGISRLNRTATFTFIED 84

```
QY 87 SMNALAPIRISLP-----VLHSDTAKKIPK--QNYHLN-----DWMEER 125
Db 85 EIYRMTESKEALPGVGFELNG-----KVPLYIKNTKLNTRCCVDFVPLIHEHNDE 140
QY 126 YRHIPGHVFTGYPCSWTYHHLPRLKEFTLHDHVREAAQFLRGLRVNGSQP----- 181
Db 141 YLHLDGRFYQ-----AWKYPSPMRNELI-----GYLKTSENFGLSPKRSNE 180
QY 182 STFYG-VHVREGDVHVMPNWKGV-ADRGYLEKALDMF--RARYSSPVFVTSNG--- 234
Db 181 TSFVTCVHIIRGDKFRV-----GFAESDENFIRKAKEFVENKATSKPCHSVVLFQDDL 234
QY 235 --MAWCREN-----INASRGDVVFAAGNGTEGSPAKDFALLTQCNTIMTI--GTGIIW 283
Db 235 PFMKNLYENESTHVSNSPDDLVYAKN-----NCDIVLITAPHSFTGWW 280
QY 284 AAYLAGDGTIVLANY-TLPSPFTLK-VFKPEAAFLPEW 319
Db 281 MGFTSGDKVYMDIRETRONVYRNGNLNPDYLYLPHW 318

RESULT 12
T20372
hypothetical protein F08A8.5 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C/Accession: T20372
R/Harris, B.
submitted to the EMBL Data Library, October 1997
A/Reference number: Z19294
A/Accession: T20372
A/Status: preliminary; translated from GE/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-383 <NUL>
A/Cross-references: EMBL:Z99710; PIDN: CAB16868.1; GSPDB: GN00019; CESP: F08A8.5
A/Experimental source: Clone F08A8
C/Genetics:
A/Gene: CESP: F08A8.5
A/Map position: 1
A/Introns: 16/1; 51/3; 77/1; 131/3; 170/3; 209/2; 244/1; 269/1; 297/3; 324/1; 354/3
C/Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 7.3%; Score 139.5; DB 2; Length 383;
Best Local Similarity 21.7%; Pred. No. 0.0061;
Matches 81; Conservative 52; Mismatches 135; Indels 105; Gaps 16;

QY 26 QMSGNTSEPMRRDSCQHNGE--LRGMFTINSIGRLGNMGYATLALARNGLRAF 83
Db 39 QKNSLSPKCEINNSQSPNYEKTLOMLFAFPSPGGIGNKLFELIHLGIATSLQRKAV 98
QY 84 IPA-----SMNALAPIRISLPVLHSD-TAKKIP--WQNYHLNDW----- 121
Db 99 INATNPSEFTLNRIQPLF-----PKLADFTLRIIPDSLVTHOOTNWGRCCVDDPSRF 154
QY 122 --MBERYRHIPGHVFTGYPCSWTYHHLPRLKEFTLHDHVRE-EAQAFLRGLRVNG 178
Db 155 LNRSDQNLILDGHFYQ-----SPKYFHHIRPQ-----VREWLAPSKLQAMRAEI 198
QY 179 SQPSTF-----VGVHVRGRGVHVMPNWKGVADRGYLEKALDMFARVSSPVFVTSN 233
Db 199 LLPAKFRDDELICHVARGFQYDGLRPSDATPTRAATFDVLYRKSHRVRVNVVLGN 258
QY 234 GMWCRENINASRGDVVFAAGNGIEGSPAK-----DF-----ALLIQ-- 269
Db 259 -----DIHATVTFEDRVAHFTFLQKPNVNSVDYSLPDISPSYTAILPTL 304
QY 270 -----CNHTIMTI--GTGIIWAAYLAGD--TIYLVLTLPDPSPLKVKFPEAA 314
Db 305 TPEITLAFSRLEFCOVLITAPSSFTGWSLVLAKRTATTYRIDLESKDGVAGEMHPEDF 364
QY 315 FLPEWGIAPDLS 327
Db 365 YPPEWIKLKTDLN 377
```

## RESULT 13

T22068

hypothetical protein F41D3.6 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C/Accession: T22068

R/White, S.

submitted to the EMBL Data Library, November 1996

A/Reference number: Z19508

A/Accession: T22068

A/Status: preliminary; translated from GE/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-500 &lt;NUL&gt;

A/Cross-references: EMBL:Z81537; PIDN: CAB04377.1; GSPDB: GN00019; CESP: F41D3.6

A/Experimental source: clone F41D3

C/Genetics:

A/Gene: CESP: F41D3.6

A/Map position: 1

A/Introns: 23/1; 59/1; 112/3; 150/3; 222/1; 278/1; 308/3; 326/3; 421/3

## Query Match

7.2%; Score 138.5; DB 2; Length 500;

Best Local Similarity 23.4%; Pred. No. 0.001;

Matches 74; Conservative 105; Mismatches 105; Indels 89; Gaps 18;

QY 52 MFTINSIGRLGNMGYATLALARNMGRIAPFASMINALAPIR----- 97

Db 49 MTTIVYKGGGNGQLFEVLSLGLGIARKLKRIA-----VFNSSDPVQLQSNLEFLNQLPRIS 103

QY 98 ---ISLPVLHSDTAK-----KIPWNYHLNDWMEERYRHIPGHVFTGYPCSWTFY 146

Db 104 EQVTSVPIEPSETIRFAISSDCCRYEISDNHLAD--ESKFLVIEGHIFQ-----SKAYF 155

QY 147 HHLRPELKEFTLHDHVREAAQFLRGLRVNGSQPSTFVGVHVRGRGVY-----HVMPNW 202

Db 156 ADMKLSI-KEWL----KPEDPEKFRMMISKTESORHK-TCVHVRGDFLTDEQH----- 203

QY 203 KGVVADRGYLEKALDMFARVSSPVFVTSNGMAWCRENINASRGDVVFAAGNGIEGSPAK 262

Db 204 --AGTDSNTYTISAIDHLRSLYHGVIFIM-SNDPKWKVHI-ADHLDY-----QK 248

QY 263 DFALL-----TQCNTIMTI--GTGIIWAAYLAGDGT--IYLANY-TLPDS 303

Db 249 DIRMKILMEDAIDDLHFSQIYCDVLITAPSSFTGWSLQSNQSAVYTRDIRETKQD 308

QY 304 PFLKVFKEPAEAFLEPW 319

Db 309 VOLQWTK-EDFYPTW 323

## RESULT 14

H71976

probable alpha(1,2)fucosyltransferase - Helicobacter pylori (strain J99)

C/Species: Helicobacter pylori

A/Variety: strain J99

C/Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 28-Jul-2000

C/Accession: H71976

R/Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.

Nature 397, 176-180, 1999

A/Title: Genomic sequence comparison of two unrelated isolates of the human gastric p

A/Reference number: A71800; MUID: 99120557; PMID: 9923682

A/Accession: H71976

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-299 &lt;ARN&gt;

A/Cross-references: GB:AE001447; GB:AE001439; NID: g4154583; PIDN: AAD05659.1; PID: g415

A/Experimental source: strain J99

C/Genetics:

A/Gene: jhp0086

C/Superfamily: Vibrio cholerae hypothetical protein wbla





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:53:11 ; Search time 10.478 seconds  
(without alignments)  
1417.108 Million cell updates/sec

Title: US-10-040-863-8\_COPY\_23\_380

Perfect score: 1921

Sequence: 1 STIIHQRIKVLQPLSEKE.....RSHFLKAKGVTCYVAGRAF 358

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	1451	75.5	1 FUT2_PIG	Q10982 sus scrofa
2	1414	73.6	1 FUT2_BOVIN	Q28113 bos taurus
3	1407.5	73.3	1 FUT2_HUMAN	Q10981 h galactosi
4	1338.5	69.7	1 FUT2_MOUSE	P97353 mus musculu
5	1296.5	67.5	1 FUT2_RABIT	Q10983 oryctolagus
6	1105.5	57.5	1 FUT1_MOUSE	Q09160 mus musculu
7	1101.5	57.3	1 FUT1_RAT	Q10980 rattus norv
8	1086	56.5	1 FUT1_RABIT	Q10979 oryctolagus
9	1051.5	54.7	1 FUT1_HUMAN	P19526 homo sapien
10	1050.5	54.7	1 FUT1_PIG	Q29043 sus scrofa
11	846	44.0	1 FUT2_RAT	Q10984 rattus norv
12	168	8.7	1 YK07_CAEEL	P34302 caenorhabdi
13	91	4.7	1 VG1D_HSVEX	P22484 equine herp
14	89	4.6	1 FMA3_PORGI	Q51826 porphyromon
15	89	4.6	1 CDH1_PHACH	Q01738 phanerochaet
16	89	4.6	1 MY10_HUMAN	Q9hd67 homo sapien
17	88.5	4.6	1 PTN7_RAT	P49445 rattus norv
18	88.5	4.6	1 VG1D_HVEEA	P24872 equine herp
19	88.5	4.6	1 VG1D_HSVEB	P24379 equine herp
20	88	4.6	1 GRB2_MOUSE	Q60631 mus musculu
21	87	4.5	1 GRB2_HUMAN	P29354 homo sapien
22	87	4.5	1 NQ01_TETH	Q56222 thermus the
23	87	4.5	1 QUTD_EMENI	P15325 emericella
24	86.5	4.5	1 CN3B_HUMAN	Q13370 homo sapien
25	85.5	4.5	1 P103_HUMAN	Q60568 homo sapien
26	85.5	4.5	1 LPH_RAI	Q02401 rattus norv
27	84.5	4.4	1 RIBB_HELPY	C25484 helicobacte
28	84.5	4.4	1 AAT_THEAO	Q33822 thermus aqu
29	84.5	4.4	1 CTPC_MYCTU	P96875 mycobacteri
30	84	4.4	1 POLG_MVEV	P03769 m genome po
31	83.5	4.3	1 TCLA_CAEEL	P03934 caenorhabdi
32	83.5	4.3	1 FMB2_PORGI	Q51825 porphyromon
33	83.5	4.3	1 CYB_PARDE	P05418 paracoccus

RESULT 1  
FUT2\_PIG  
ID FUT2\_PIG STANDARD; PRT; 340 AA.  
AC Q10982: Q29044; Q19100;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)Ft 2)  
DE D-galactoside 2-alpha-L-fucosyltransferase 2.  
DE (Fucosyltransferase 2).  
GN FUT2.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97468270; PubMed=9321466;  
RA Meiferink E., Fries R., Voegeli P., Masabanda J., Wigger G., Stricker C., Neuenschwander S., Bertschinger H.U., Stranzinger G.; "Two alpha(1,2) fucosyltransferase genes on porcine chromosome 6q11 are closely linked to the blood group inhibitor (S) and Escherichia coli F18 receptor (ECP18R) loci."; Mamm. Genome 8:736-741(1997).  
RL [2]  
RN SEQUENCE FROM N.A.  
RA Cohnsey S., Mouhtouris E., McKenzie I.F.C., Sandrin M.S.; "Molecular cloning and characterization of the pig secretor type alpha(1,2)fucosyltransferase."; alpha(1,2)fucosyltransferase."; Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE OF 69-334 FROM N.A.  
RA Petit J.M.; Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE OF 69-80; 119-133 AND 316-334.  
RC TISSUE=Submaxillary gland;  
RX MEDLINE=96064742; PubMed=7592879;  
RA Thurlin J., Blaszczyk-Thurin M.; "Porcine submaxillary gland GDP-L-fucose: beta-D-galactoside alpha-2-L-fucosyltransferase is likely a counterpart of the human Secretor gene-encoded blood group transferase."; J. Biol. Chem. 270:26577-26580(1995).  
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R + alpha-L-fucosyl-1,2-beta-D-galactosyl-R.  
CC -!- PATHWAY: Glycosylation.  
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

Q61409 mus musculu  
P76484 escherichia  
P45784 vibrio chol  
P72933 synecocyst  
P87503 human adeno  
P97846 rattus norv  
P78357 homo sapien  
P87379 xenopus lae  
Q92140 helicobacte  
P27339 rattus norv  
P47527 mycoplasma  
O54991 mus musculu

## ALIGNMENTS







Db 261 NIDASRGDVVFAGNGLESPAKDFALLTQCNHTYMT-GTGFNAAYLTGGDTVYLANYTA 320

QY 301 PDSFELKVFPEAFIPENTWGIAPDLSPL 330

Db 321 PDSFELKVFPEAFIPENTWGIAPDLSPL 350

RESULT 6

FUT1\_MOUSE

ID FUT1\_MOUSE STANDARD; PRT; 376 AA.

AC 009160;

DT 15-DEC-1998 (Rel. 37, Created);

DT 15-DEC-1998 (Rel. 37, Last sequence update);

DT 15-JUN-2002 (Rel. 41, Last annotation update);

DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)

DE (Fucosyltransferase 1).

GN FUT1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NIH Swiss;

RX MEDLINE=97454449; PubMed=9355741;

RA Domino S.E., Hiraiwa N., Lowe J.B.;

RT "Molecular cloning, chromosomal assignment and tissue-specific expression of a murine alpha(1,2)fucosyltransferase expressed in thymic and epidermal epithelial cells.;"

RL Biochem. J. 327:105-115(1997).

CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP + alpha-L-fucosyl-1,2-beta-D-galactosyl-R.

CC -!- PATHWAY: Glycosylation.

CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND FORM IN TRANS CISTERNAE OF GOLGI.

CC -!- TISSUE SPECIFICITY: IN THE ADULT, HIGHLY EXPRESSED IN PANCREAS, TESTIS AND EPIDIDYMS AND TO A LESSER EXTENT IN THYMUS, LUNG, STOMACH, SMALL INTESTINE, COLON, SPLEEN AND UTERUS. NOT EXPRESSED IN BRAIN, HEART, SKELETAL MUSCLE, KIDNEY, LIVER AND BONE MARROW.

CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE GALACTOSIDE 2-L-FUCOSYLTRANSFERASE.

CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U90553; AAC53492.1; -

DR MGB; 109375; Fut1.

DR Interpro; IPR002516; GP\_11.

DR Pfam; PF01531; Glyco-transf 11; 1.

DR

KW Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane; Signal-anchor; Golgi stack.

KW DOMAIN 1

FT TRANSMEM 9 26 CYTOPLASMIC (POTENTIAL).

FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).

FT DOMAIN 27 376 LUMENAL, CATALYTIC (POTENTIAL).

FT CARBOHYD 67 67 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 302 302 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 328 328 N-LINKED (GLCNAC...) (POTENTIAL).

SQ SEQUENCE 376 AA; 42255 MW; 21FD24CFE2C4106 CRC64;

Query Match 57.5%; Score 1105.5; DB 1; Length 376;

Best Local Similarity 62.8%; Pred. No. 3.6e-86;

Matches 208; Conservative 38; Mismatches 76; Indels 9; Gaps 2;

QY 15 PLSEKEPMTOMSSGNTSEPMRDSQHGSLRGMTINSIGRGNGMGVATLAL 74

Db 53 PVAMVCLPFLPLOTSGNSGPSCEQ-----SSLSGTVITPGRGNGMGVATLAL 104

QY 75 ARMGRLAFIPASMHNALAPIERISLPVLHSDTAKKIPQWYHLNDWMEERYRHPGHV 134

Db 105 AOLNGRAFIQEMHAALAPVFRISLPVLDEVDLSLPWQHLVLDHNMSEYSHLEDPL 164

QY 135 RFTGPGSWTFYHHLRPEILKEFTLHDHVREREAQAFRLGRLVNGSQPS-TFVGTVVRGD 193

Db 165 KLSGFPSCSWTFHHLRPEILKEFTLHDHVREREAQAFRLGRLVNGSQPS-TFVGTVVRGD 224

QY 194 YVHVPNVWGVVADRGYLEKALDMFRARYSSPVFVVISNGMAWCRNINASRGDGVVFG 253

Db 225 YLEVPNVRWGVGDRAVLOQAMDWRARHKDPFVTVTSNGMKWCLENIDSHGDVVFG 284

QY 254 NGIEGSPAKDFALLTQCNHTIMTIGTGTGWAAYLAGDITVLANYTLPDSFFLKVFPEA 313

Db 285 NGGEGIPGKDFALLTQCNHTIMTIGTGTGWAAYLAGDITVLANYTLPDSFFLKVFPEA 344

QY 314 AFLPEWVGIPADLSPLLKALTPACPRSHPHL 344

Db 345 AFLPEWVGINADLSPLQAQDFPKWPSLFL 375

RESULT 7

FUT1\_RAT

ID FUT1\_RAT STANDARD; PRT; 376 AA.

AC Q10980;

DT 01-OCT-1996 (Rel. 34, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)

DE (Fucosyltransferase 1).

GN FUT1 OR FTA.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-Colon cancer;

RA Soejima M., Wang B., Koda Y., Kimura H.;

RT "Two distinct rat GDP-L-fucose:b-D-galactoside 2-alpha-L-fucosyltransferase genes.;"

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 169-310 FROM N.A.

RC STRAIN=BDIX;

RX MEDLINE=94280382; PubMed=8010942;

RA Piau J.-P., Labarriere N., Dabouis G., Denis M.G.;

RT "Evidence for two distinct alpha(1,2)-fucosyltransferase genes differentially expressed throughout the rat colon.;"

RL Biochem. J. 300:623-626(1994).

CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP + alpha-L-fucosyl-1,2-beta-D-galactosyl-R.

CC -!- PATHWAY: Glycosylation.

CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).

CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.

CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AB015637; BAA31130.1; -

DR EMBL; AB006137; BAA21741.1; -

```

DR EMBL: L26009; AAB41514.1; -.
DR InterPro: IPR002516; GT 11.
DR Pfam: PF01531; Glyco_transf_11; 1.
DR Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 13 29 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT (POTENTIAL).
FT DOMAIN 30 376 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 376 AA; 42416 MW; C350C737C758B7F8 CRC64;

Query Match 57.3%; Score 1101.5; DB 1; Length 376;
Best Local Similarity 62.3%; Pred. No. 7.9e-86;
Matches 207; Conservative 37; Mismatches 77; Indels 11; Gaps 2;

QY 15 PLSEKELPMTOMSSNGTESPERDRSEHQHNGELRGMTINSIGRLGNOMGEYATILFAL 74
DB 53 PVAMVCLPYSPNASSGSPSCPE-----QSLLSGTWITPGRGNGMGQYATLLAL 103
QY 75 ARNGLAFTPASWHAHALAPIFRISLPVLSHDTAKKIPWNYHLNDWMEERYRHIPGHEV 134
DB 104 AQLNGRAFTQPEMHTTAPVFRISLPVLPDPEVDSLTPWQHLVLDWMSSEYSHLEDPEL 163
QY 135 RFTGPCSWTFYHHLRPEILKEFTLHDVREPAQAFLGLRVN--GSQSPTFVGVHVRG 192
DB 164 KLSGFCFSWTFHHLREQINREFTLHDHLDREDAQRLLSLGLRPGAGIRPIRYVGVHVRG 223
QY 193 DYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVVT-SNGMAWCRENINASRGDVVFA 252
DB 224 DYLEVMPNRKGVVGDVAYLQKAMDWFRARHKDPIFVVISNGMWGLENDLSHGQDVVFA 283
QY 253 GNGIESPAKDFALLTQCNTHTMTIGTFEINAYLAGGDTIYIANYTLDPSPFLKYFKPE 312
DB 284 GNGQESTPKDFALLTQCNTHTMTIGTFEINAYLAGGDTIYIANYTLDPSEFLKIFRK 343
QY 313 AAFLEPWVGIPADLSPLLKALTPACPRSHFL 344
DB 344 AAFLEPWVGINADLSPLQAFQDFWETDSFLRL 375

RESULT 8
FUT1_RABIT
ID FUT1_RABIT STANDARD; PRT; 373 AA.
AC Q10979;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)
DE D-galactoside 2-alpha-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)
DE FUT1 OR RFT-I.
GN Oryctolagus cuniculus (Rabbit).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95238380; PubMed=7721792;
RA Hitoshi S., Kusunoki S., Kanazawa I., Tsuji S.;
RT "Molecular cloning and expression of two types of rabbit beta-galactoside alpha 1,2-fucosyltransferase.";
RL J. Biol. Chem. 270:8844-8850(1995).
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP + alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND FORM IN TRANS CISTERNAE OF GOLGI.
CC -!- TISSUE SPECIFICITY: BRAIN.
CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE

```

```

CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
CC TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF
CC MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: X80226; CAA56513.1; -.
DR InterPro: IPR002516; GT 11.
DR Pfam: PF01531; Glyco_transf_11; 1.
DR Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 13 29 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT (POTENTIAL).
FT DOMAIN 30 373 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 373 AA; 42098 MW; 0A47A1786231525C CRC64;

Query Match 56.5%; Score 1086; DB 1; Length 373;
Best Local Similarity 58.7%; Pred. No. 1.6e-84;
Matches 206; Conservative 50; Mismatches 59; Indels 36; Gaps 5;

QY 1 STTIHQQRIVK-----LOPLSEK-----LPMITQMS-----SGNTSPEMRRD 40
DB 23 SFLLHLHQDLFRNGLALSPLCLERQVPAPVAIVCLFVTPSNASSCAGRPAAPS---- 78
QY 41 SEQHNGELRGMTINSIGRLGNOMGEYATILFALARMNGRLAPFAPSMNALAPIFRISL 100
DB 79 -----GIWTHPDGRFGNQMGQYATLLALQNLGRRAFLPAMHAALAPVERITL 128
QY 101 PVLHSDTAKKIPWNYHLNDWMEERYRHIPGVHVRGTYVHVPNVKGVVADRGYLEKALDM 160
DB 129 PVLAPENVNRTSWKQLLHDWSEYSRIEDPEFLKFGPCSWTFHHLRPEILKEFTLH 188
QY 161 DHVREAAQFLRGLRV--NGSQPSIFVGVHVRGTYVHVPNVKGVVADRGYLEKALDM 218
DB 189 DHLREAAQLLGLRGTGARPTFVGHVRRGDYLVQMPQRKGVGDRAYLQQAMDW 248
QY 219 FRARYSSPVFVTSNGMAWCRENINASRGDVVFPAGNIEGSPAKDFALLTQCNTHTMTIG 278
DB 249 FRARHAPIFVVTSGMKWCWENIDASRGDVVFPAGNIEGSPAKDFALLTQCNTHTMTIG 308
QY 279 TFSIWAAYLAGGDTIYIANYTLDPSPFLKVFKEPAFLPEWVGIPADLSPL 329
DB 309 TEGFWAAVLAGGDTIYIANYTLDPSEFLKIFKPEAFIPEWVGINADLSPV 359

RESULT 9
FUT1_HUMAN
ID FUT1_HUMAN STANDARD; PRT; 365 AA.
AC P19526;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)
DE D-galactoside 2-alpha-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)
DE FUT1.
DE Homo sapiens (Human).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```







```
YK07_CAEEL
ID YK07_CAEEL STANDARD; PRT; 365 AA.
AC P34302;
DT 01-FEB-1994 (Rel. 28, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative glycosyl transferase C06E1.7 in chromosome III (EC 2.-.-.-).
GN C06E1.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M., Coulson A.,
RA Craxton J., Burton J., Connell M., Copsey T., Cooper J., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smaldon N., Smith A., Smith K., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP REVISIONS.
RA Waterston R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L16559; AAA27932.2;
CC WormPep; C06E1.7; CE30483.
CC InterPro; IPR002516; GI.11.
CC Pfam; PF01531; Glyco.transf.11; 1.
CC Hypothetical protein; Transferase; Glycosyltransferase.
CC SEQUENCE 365 AA; 41991 MW; B5FBCA363F51977F CRC64;
CC
CC Query Match 8.7%; Score 168; DB 1; Length 365;
CC Best Local Similarity 22.0%; Pred. No. 7.2e-07;
CC Matches 71; Conservative 48; Mismatches 115; Indels 88; Gaps 16;
CC
CC YQ 57 SIGRLGNMGEXAIFALARMNGR--LATIPASMHNALAPIFRISLPVLHSD----- 106
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC Db 70 STARLANHFIELVSVTGMKSLNRKPAIFVEDSKSYNLLITGVKVLPLGLDEQIFEPV 129
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC YQ 107 --TAKKIP-----WQNYHLNDWMEERYRHIPGHEVRTGYPCSWTFYHLLRPEILKEF 157
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC Db 130 HNKATKVPILSEKCCIFDNPDKFNNSISVILHUGHYQ-----SWKYF----- 172
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC YQ 158 TLHDHVRERAQAFRLGL-----RVNGSQPTEFG-----VHVRGDIY---HVMNVNWKGVVA 207
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC Db 173 ---DKYKEKQVSKPAIDFSLPNSDSSNFISRICIHRIRDFDVGQHSNV----- 223
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC YQ 208 DRGLEKALDMFRAR-----YSSPVFVTSNGMAMCWMCRINASRGD 248
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC Db 224 --SFTKPALEFKEREQKDVNKKMLTVIMGDDPDFEAKMF-----EGTVRAKKEAKEET 277
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC YQ 249 VVFAGNGIGSGFAKDFALL-TCQNTHTIMI--GTGCIWAIYLAGDGTIYLANI-LPDPSP 304
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
```

```
Db 278 KYFVS---ENTPQDDLAYSHYSCDATTLPSTFGWVLGYLSKGOAVYQDIRSTNDVN 334
QY 305 FLK-VFKPEAAFLPEWVGIPAD 325
: | | | | : | | | |
Db 335 YKKGLVDFDDFEVPSWTSIMLD 356
: | | | | : | | | |
RESULT 13
VGLD_HSVK
ID VGLD_HSVK STANDARD; PRT; 442 AA.
AC P22484;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein D precursor (glycoprotein 17/18).
GN GD OR GP17/18 OR 72.
OS Equine herpesvirus type 1 (strain Kentucky A) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91082407; PubMed=1845821;
RA Flowers C.C., Eastman E.M., O'Callaghan D.J.;
RT "Sequence analysis of a glycoprotein D gene homolog within the unique
RT short segment of the EHV-1 genome.";
RL Virology 180:175-184(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92263758; PubMed=1315673;
RA Colle C.F., III, Flowers C.C., O'Callaghan D.J.;
RT "Open reading frames encoding a protein-kinase, homolog of
RT glycoprotein gX of pseudorabies virus, and a novel glycoprotein map
RT within the unique short segment of equine herpesvirus type 1.";
RL Virology 188:545-557(1992).
```

```
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN D FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M62923; AAA46081.1;
CC EMBL: M86931; -; NOT_ANNOTATED_CDS.
CC EMBL: M87497; AAA46073.1; ALT_INIT.
CC PIR: A38518; VGBEPA.
CC InterPro; IPR002896; Herpes_glycop_D.
CC Pfam; PF01537; Herpes_glycop_D; 1.
CC Glycoprotein; Signal; Transmembrane.
CC SIGNAL 1 19 POTENTIAL.
CC CHAIN 20 442 GLYCOPROTEIN D.
CC DOMAIN 20 405 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 406 422 POTENTIAL.
CC DOMAIN 423 442 CYTOPLASMIC (POTENTIAL).
CC CARBOHYD 103 103 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 111 111 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 347 347 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 396 396 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 442 AA; 49908 MW; 323CDDCA9C9762F05 CRC64;
```

```
Query Match 4.7%; Score 91; DB 1; Length 442;
Best Local Similarity 22.7%; Pred. No. 3;
Matches 54; Conservative 31; Mismatches 103; Indels 50; Gaps 10;
CC
CC YQ 6 LOORIVKLO-----PLSEKELPMTQSSNGTESPEMRDSEQHNGELRGFTI 55
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC Db 234 LYRVEIDGRRIYDFSVTIPSEPCIAFELNFGN---PDRCKTPEQYSRG- --VF-R 287
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC YQ 56 SIEGLNSGMGEVAT--LEFALRMNGRLAIPASMHNALAPIFRISLPVLHSDTAKKI 113
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
```

Db 288 RFLGEFNPGEHMTWKRFVYDGNL-----PVQFYFAQA-----FARVPPP 331

QY 114 ONYHLNDWMEERYH-----IPGHVFRTGYPCSWTFYHHLRPEILKEFTLHDHVRDEAQ 168

Db 332 DNHFGFDSVESEITQNTDKPKQADPKPQPKWPSIKHLVPL-----DEVDEVIE 384

QY 169 AFLRGLVNGSQSTFGVGVHVRGDDYVHVMPNVKGVVADRGKYLEKALDMFRARYSSP 226

Db 385 PVTKPPKTSKN-STFGVIGSVGLGIAGLVGLVGLTYCLRR--KKELKVCTERLDSF 438

RESULT 14

FM33 PORGI

ID FNA3\_PORGI STANDARD; PRT; 353 AA.

AC G51826; 10

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE Major fimbrial subunit protein, type III precursor (Fimbrillin)

DE (Fimbrillin).

GN FIMA.

OS Porphyromonas gingivalis (Bacteroides gingivalis).

OC Bacteria; Bacteroidetes; Bacteroides; Bacteroidales;

OC Porphyromonadaceae; Porphyromonas.

OX NCBI\_TaxID=837;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=6/26;

RX MEDLINE=94071950; PubMed=7902712;

RA Fujiwara T., Morishima S., Takahashi I., Hamada S.;

RT "Molecular cloning and sequencing of the fimbrial gene of

RT Porphyromonas gingivalis strains and characterization of recombinant

RT proteins";

RL Biochem. Biophys. Res. Commun. 197:241-247(1993).

RN [2]

RP FUNCTION, AND CLASSIFICATION INTO TYPES.

RX MEDLINE=21614934; PubMed=11748193;

RA Nakagawa I., Amano A., Kubonawa M., Nakamura T., Kawabata S.,

RA Hamada S.;

RT "Functional differences among FimA variants of Porphyromonas

RT gingivalis and their effects on adhesion to and invasion of human

RT epithelial cells.";

RL Infect. Immun. 70:277-285(2002).

CC -!- FUNCTION: Fimbrillin is the structural subunit of the fimbriae,

CC that are filamentous appendages on the cell surface. Fimbriae of

CC P. gingivalis are recognized as a major virulence factor as they

CC mediate cell adhesion and play an important role in invasion of

CC periodontal tissues.

CC -!- SUBCELLULAR LOCATION: Fimbria.

CC -!- SIMILARITY: BELONGS TO THE P.GINGIVALIS FIMBRILLIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; D17801; BAA04627.1; -

FW Fimbria; Virulence.

KW PROPEP 1 10 BY SIMILARITY.

FT CHAIN 11 353 MAJOR FIMBRILL SUBUNIT PROTEIN, TYPE III.

SQ SEQUENCE 353 AA; 38024 MW; 7FBE4FBF42FEA2AB CRC64;

Query Match 4.6%; Score 89; DB 1; Length 353;

Best Local Similarity 22.1%; Pred. No. 3.3;

Matches 64; Conservative 38; Mismatches 110; Indels 78; Gaps 12;

QY 23 MTIQSSCNTSPSPRRDSEQ-----HGNGELRGMTTINSIGRLNQMG---EYATLFA 73

Db 80 LTTLETGNGPFAAGLIMTAEPVEVTLVAGN-----NYYGVDGSGGNGISQGTPLF 130

QY 74 LARNGRIAP--IPASHNRLAPIFRISLPVLHSDTAKK----- 110

Db 131 IKRVHARIATKIEVMTSQSYANKYNFAPENIYALVAKKSNLFGASLANSDDAYLTGSL 190

QY 111 -----IPWNYHLNDWMEERYH-----PGHFVFTGYPCSWTFYHH--LRPEIL- 154

Db 191 ITFNGAYSPANYTHVDWLGRDYTEIGATVNTPKGF-----YVLESTYAGNAGLRPTILC 245

QY 155 --KEFTLHDHV-----REEAQFLGLRVNGSQSTFG--VGHVHVRGDDYVHVMPNVKGV 206

Db 246 VKGLTKHDKGTALSSSEMTAFAFNAGWIVANNDPTYYPLVNPESNNYTYTGEAVEGKI 305

QY 207 ADRGYLEKALDMFRARYSSPVFVVTNSNGMAWCENINASRGDVVFPAGNI 256

Db 306 VRNHKFDINLTITGPGTNNPENPIT-----ESANLNYCVVAARKGV 347

RESULT 15

CDH\_PHACH

ID CDH\_PHACH STANDARD; PRT; 773 AA.

AC Q01738; C00047;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cellobiose dehydrogenase precursor (EC 1.1.5.1) (CDH) (Cellobiose-

DE quinone oxidoreductase).

GN CDH-1 AND CDH-2.

OS Phanerochaete chrysosporium.

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;

OC Aphyllophorales; Corticiaceae; Phanerochaete.

OX NCBI\_TaxID=5306;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OGC101;

RX MEDLINE=97077226; PubMed=8919793;

RA Li B., Nagalla S.R., Renganathan V.;

RT "Cloning of a cDNA encoding cellobiose dehydrogenase, a

RT hemoflavoenzyme from Phanerochaete chrysosporium.";

RL Appl. Environ. Microbiol. 62:1329-1335(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=OGC101;

RX MEDLINE=97176414; PubMed=9023960;

RA Li B., Nagalla S.R., Renganathan V.;

RT "Cellobiose dehydrogenase from Phanerochaete chrysosporium is encoded

RT by two allelic variants.";

RL Appl. Environ. Microbiol. 63:796-799(1997).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 19-208.

RX MEDLINE=20139694; PubMed=10673428;

RA Hallberg B.M., Bergfors T., Boeckro K., Pettersson G., Henriksson G.,

RA Dyne C.;

RT "A new scaffold for binding haem in the cytochrome domain of the

RT extracellular flavocytochrome cellobiose dehydrogenase.";

RL Structure 8:79-88(2000).

CC -!- FUNCTION: DEGRADATES BOTH LIGNIN AND CELLULOSE. OXIDIZES CELLOBIOSE

CC TO CELLOBIONOLACTONE.

CC -!- CATALYTIC ACTIVITY: Cellobiose + a quinone = cellobiono-1,5-

CC lactone + a phenol.

CC -!- COFACTOR: ONE FAD AND ONE HEME B.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE GMC

CC OXIDOREDUCTASES FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U46081; AAC49277.1; -

Search completed: May 27, 2003, 15:07:13  
Job time : 11.478 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	2032	100.0		380	11	O9JK44	O9JK44 rattus nor
2	1896	93.3		353	11	O55025	O55025 rattus nor
3	1883	92.7		354	11	O35087	O35087 rattus nor
4	1878	92.4		354	11	O3R275	O3R275 rattus nor
5	1674.5	82.4		347	11	O9J227	O9J227 mus musculu
6	1670.5	82.2		347	11	O70504	O70504 mus musculu
7	1547.5	76.2		321	11	O920W2	O920W2 mus spicile
8	1547.5	76.2		321	11	O91VFO	O91VFO mus musculu
9	1532.5	75.4		321	11	O920W3	O920W3 mus musculu
10	1499.5	73.8		343	6	O77486	O77486 gorilla gor
11	1495.5	73.6		343	6	O77485	O77485 pan troglod
12	1495.5	73.6		343	6	O9TUD3	O9TUD3 gorilla gor
13	1488.5	73.3		338	6	O9TUD5	O9TUD5 pan troglod
14	1488.5	73.3		347	6	O29505	O29505 oryctolagus
15	1487.5	73.2		343	6	O77487	O77487 ponga pygma
16	1478.5	72.8		343	6	O9TTC7	O9TTC7 hylabates l

Ov 121 SE.PVT.HSDTAKKT

.....

100

[illegible]

```
Db 121 SLPVLHSDTAKKIPWQNYHLNDWMEERYHIFGHFVRFYPCSWTFYHHLRPEILKEPT 180
QY 181 LHDHVREEAQAFRLGLRVNGSQPSTFVGHVHVRGDYVHVMPNWKGVVADRGYLEKALDM 240
Db 181 LHDHVREEAQAFRLGLRVNGSQPSTFVGHVHVRGDYVHVMPNWKGVVADRGYLEKALDM 240
QY 241 FRARYSSPVVVTSGMAMCARENINASRGDVVFAGNGIEGSPAKDFALLTQCNHTIMTIG 300
Db 241 FRARYSSPVVVTSGMAMCARENINASRGDVVFAGNGIEGSPAKDFALLTQCNHTIMTIG 300
QY 301 TFGIWAAYLAGGDTIYLANITLPDPSFLKVPKPEAAFLPEWVGIPADLSPLKALTPACP 360
Db 301 TFGIWAAYLAGGDTIYLANITLPDPSFLKVPKPEAAFLPEWVGIPADLSPLKALTPACP 360
QY 361 RSHFHLKAGVTCYVAGRAF 380
Db 361 RSHFHLKAGVTCYVAGRAF 380
RESULT 2
Q55025
AC O55025 PRELIMINARY; PRT; 353 AA.
DT 01-JUN-1998 (TREMELrel. 06, Created)
DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE Alpha 1,2 fucosyltransferase (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BUFFALO RAT;
RX MEDLINE=98342056; PubMed=9675030;
RA Sherwood A.L., Holmes E.H.;
RT "Cloning and expression of the catalytic domain from rat hepatoma H35
RT cell GDP-fucose: GMI alpha 1-->2fucosyltransferase, an enzyme which is
RT activated during early stages of chemical carcinogenesis in rat
RT liver.";
RL Arch. Biochem. Biophys. 355:215-221(1998).
DR EMBL: AF042743; AAC14695.1;
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1
SQ SEQUENCE 353 AA; 39738 MW; 609D64EB222C9585 CRC64;
Query Match 93.3%; Score 1896; DB 11; Length 353;
Best Local Similarity 100.0%; Pred. No. 3.8e-164;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 28 LQORIVKLOPLSEKELPMTQSSGNTSPENRSEQHGNGELRGMTFINSIGRLGNOM 87
Db 1 LQORIVKLOPLSEKELPMTQSSGNTSPENRSEQHGNGELRGMTFINSIGRLGNOM 60
QY 88 GYATLALARNMGRALFIPASMHNALAPIFRISLFLVHSDTAKKIPWQNYHLNDWMEER 147
Db 61 GYATLALARNMGRALFIPASMHNALAPIFRISLFLVHSDTAKKIPWQNYHLNDWMEER 120
QY 148 YRHIPGHFVRFTGYPCSWTFYHHLRPEILKEPTLHDHVREEAQAFRLGLRVNGSQPSTFV 207
Db 121 YRHIPGHFVRFTGYPCSWTFYHHLRPEILKEPTLHDHVREEAQAFRLGLRVNGSQPSTFV 180
QY 208 GYHVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVVVTSGMAMCARENINAS 267
Db 181 GYHVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVVVTSGMAMCARENINAS 240
QY 268 RGVVPFAGNGIEGSPAKDFALLTQCNHTIMTIGTFGIWAAYLAGGDTIYLANITLPDPSF 327
Db 241 RGVVPFAGNGIEGSPAKDFALLTQCNHTIMTIGTFGIWAAYLAGGDTIYLANITLPDPSF 300
```

```
QY 328 LKVFKEAAFLPEWVGIPADLSPLKALTPACPRSHFHLKAKGVTCYVAGRAF 380
Db 301 LKVFKEAAFLPEWVGIPADLSPLKALTPACPRSHFHLKAKGVTCYVAGRAF 353
RESULT 3
Q55087
ID O35087 PRELIMINARY; PRT; 354 AA.
AC O35087;
DT 01-JAN-1998 (TREMELrel. 05, Created)
DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON CANCER;
RA Soejima M., Wang B., Koda Y., Kimura H.;
RT "Two distinct rat GDP-L-fucose:b-D-galactoside 2-a-L-
RT fucosyltransferase genes.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DDJ databases.
DR EMBL: AB006138; BAA21742.1;
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 354 AA; 39995 MW; 86364444889215BA1 CRC64;
Query Match 92.7%; Score 1883; DB 11; Length 354;
Best Local Similarity 100.0%; Pred. No. 5.8e-163;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASAQVPFSPPLAHFLIFVFTVSTIIHLQORIVKLOPLSEKELPMTQSSGNTSPENR 60
Db 1 MASAQVPFSPPLAHFLIFVFTVSTIIHLQORIVKLOPLSEKELPMTQSSGNTSPENR 60
QY 61 RSEQHNGELRGMTFINSIGRLGNOMGEYATLFLALARMNGRLAFIPASMHNALAPIFR 120
Db 61 RSEQHNGELRGMTFINSIGRLGNOMGEYATLFLALARMNGRLAFIPASMHNALAPIFR 120
QY 121 SLPVLHSDTAKKIPWQNYHLNDWMEERYHIFGHFVRFYPCSWTFYHHLRPEILKEPT 180
Db 121 SLPVLHSDTAKKIPWQNYHLNDWMEERYHIFGHFVRFYPCSWTFYHHLRPEILKEPT 180
QY 181 LHDHVREEAQAFRLGLRVNGSQPSTFVGHVHVRGDYVHVMPNWKGVVADRGYLEKALDM 240
Db 181 LHDHVREEAQAFRLGLRVNGSQPSTFVGHVHVRGDYVHVMPNWKGVVADRGYLEKALDM 240
QY 241 FRARYSSPVVVTSGMAMCARENINASRGDVVFAGNGIEGSPAKDFALLTQCNHTIMTIG 300
Db 241 FRARYSSPVVVTSGMAMCARENINASRGDVVFAGNGIEGSPAKDFALLTQCNHTIMTIG 300
QY 301 TFGIWAAYLAGGDTIYLANITLPDPSFLKVPKPEAAFLPEWVGIPADLSPLK 353
Db 301 TFGIWAAYLAGGDTIYLANITLPDPSFLKVPKPEAAFLPEWVGIPADLSPLK 353
RESULT 4
Q9R275
ID Q9R275 PRELIMINARY; PRT; 354 AA.
AC Q9R275;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
```

```

[1]
RN      SEQUENCE FROM N.A.
RC      STRAIN=BDIX;
RX      MEDLINE=2109374; PubMed=11179967;
RA      Bureau V., Marionneau S., Cailleteau-Thomas A., Le Moullac-Vaidye B.,
RA      Liehr T., Le Pendu J.;
RT      "Comparison of the three rat GDP-L-fucose: beta-D-galactoside 2-
RL      Eur. J. Biochem. 268:1006-1019(2001).
DR      EMBL; AF131238; AAD24469.1; -
DR      InterPro; IPR002516; GT_11.
DR      Pfam; PF01531; Glyco_transf_11; 1.
KW      Glycosyltransferase; Transferase.
SQ      SEQUENCE 354 AA; 39983 MW; 123E8C8379E8559E CRC64;

Query Match      92.4%; Score 1878; DB 11; Length 354;
Best Local Similarity 99.7%; Pred. No. 1.7e-162;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MASAOVPSFPLAHLFIHFVFTSTIIHLQQRIVKLOPLSEKELPMTTOMSSGNTESPENR 60
Db 1 MASAOVPSFPLAHLFIHFVFTSTIIHLQQRIVKLOPLSEKELPMTTOMSSGNTESPENR 60
Qy 61 RDSQHGNGELRGMTINSIGRLGNMGGEYATLFLALARMNGRLAFIPASMHNALAPIERI 120
Db 61 RDSQHGNGELRGMTINSIGRLGNMGGEYATLFLALARMNGRLAFIPASMHNALAPIERI 120
Qy 121 SLPLVHSDTAKKIPQWNYHLNDMEERYRHIPGHVFRTGYPCSWTFYHHLRPEILKEFT 180
Db 121 SLPLVHSDTAKKIPQWNYHLNDMEERYRHIPGHVFRTGYPCSWTFYHHLRPEILKEFT 180
Qy 131 LHDHVREAAQAFRLGLRVNGSQPSTFVGHVHVRGDIYVHVPNVKGVVADRGYLEKALDM 240
Db 131 LHDHVREAAQAFRLGLRVNGSQPSTFVGHVHVRGDIYVHVPNVKGVVADRGYLEKALDM 240
Qy 241 FRARYSSPVFVVTNSGMAWCARENINASRGDVVFAGNGTEGSPAKDFALLTCQNHFTMTIG 300
Db 241 FRARYSSPVFVVTNSGMAWCARENINASRGDVVFAGNGTEGSPAKDFALLTCQNHFTMTIG 300
Qy 301 TFGIWAAYLAGDGTIYLANYLTPDSPFLKVPKPEAAFLPEWVGIPADLSPLIK 353
Db 301 TFGIWAAYLAGDGTIYLANYLTPDSPFLKVPKPEAAFLPEWVGIPADLSPLIK 353

RESULT 5
Q9JL27 PRELIMINARY; PRT; 347 AA.
ID Q9JL27
AC Q9JL27;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15; Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Alpha(1,2)fucosyltransferase FUT2 (EC 2.4.1.69).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/OLA;
RX MEDLINE=21316545; PubMed=11323419;
RA Domino S.E., Zhang L., Lowe J.B.;
RA "Molecular Cloning, Genomic Mapping, and Expression of Two Secretor
RT Blood Group alpha (1,2)fucosyltransferase Genes Differentially
RT Regulated in Mouse Uterine Epithelium and Gastrointestinal Tract.";
RL J. Biol. Chem. 276:23748-23756(2001).
DR EMBL; AF214656; AAF45146.1; -
DR MGB; MGI:109374; Fut2.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 347 AA; 39243 MW; D5F441046CC8E79 CRC64;

```

```

Query Match      82.4%; Score 1674.5; DB 11; Length 347;
Best Local Similarity 89.8%; Pred. No. 5.3e-144;
Matches 317; Conservative 12; Mismatches 17; Indels 7; Gaps 2;

Qy 1 MASAOVPSFPLAHLFIHFVFTSTIIHLQQRIVKLOPLSEKELPMTTOMSSGNTESPENR 60
Db 1 MASAOVPSFPLAHLFIHFVFTSTIIHLQQRIVKLOPLSEKEL-QAVQMSSPNAARTDMQ 59
Qy 61 RDSQHGNGELRGMTINSIGRLGNMGGEYATLFLALARMNGRLAFIPASMHNALAPIERI 120
Db 60 Q-----SAKLOGIFTINSIGRLGNMGGEYATLFLALARMNGRLAFIPESMHNALAPIERI 113
Qy 121 SLPLVHSDTAKKIPQWNYHLNDMEERYRHIPGHVFRTGYPCSWTFYHHLRPEILKEFT 180
Db 114 SLPLVHSDTARRIPQWNYHLNDMEERYRHIPQGVRTGYPCSWTFYHHLRPEILKEFT 173
Qy 131 LHDHVREAAQAFRLGLRVNGSQPSTFVGHVHVRGDIYVHVPNVKGVVADRGYLEKALDM 240
Db 174 LHDHVREAAQAFRLGLRVNGSQPSTFVGHVHVRGDIYVHVPNVKGVVADRGYLEKALOR 233
Qy 241 FRARYSSPVFVVTNSGMAWCARENINASRGDVVFAGNGTEGSPAKDFALLTCQNHFTMTIG 300
Db 234 FRARYSSPVFVVTNSGMAWCARENINTSLGDDVVFAGNGTEGSPAKDFALLTCQNHFTMTIG 293
Qy 301 TFGIWAAYLAGDGTIYLANYLTPDSPFLKVPKPEAAFLPEWVGIPADLSPLIK 353
Db 294 TFGIWAAYLAGDGTIYLANYLTPDSPFLKIPKPAALPEWVGIPADLSPLIK 346

RESULT 6
Q70504 PRELIMINARY; PRT; 347 AA.
ID Q70504
AC Q70504;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07; Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase (EC
DE 2.4.1.69).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=GASTROINTESTINAL TRACT;
RX MEDLINE=20471982; PubMed=11018479;
RA Lin B., Hayashi Y., Saito M., Sakakibara Y., Yanagisawa M.,
RA Iwamori M.;
RA "GDP-fucose: beta-galactoside alphaL,2-fucosyltransferase, MFUT-II,
RT and not MFUT-I or -III, is induced in a restricted region of the
RT digestive tract of germ-free mice by host-microbe interactions and
RT cycloheximide.";
RL Biochim. Biophys. Acta 1487:275-285(2000).
DR EMBL; AF064792; AAC16887.1; -
DR MGB; MGI:109374; Fut2.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 347 AA; 39215 MW; 616CCB11581C4179 CRC64;

Query Match      82.2%; Score 1670.5; DB 11; Length 347;
Best Local Similarity 89.5%; Pred. No. 1.2e-143;
Matches 316; Conservative 13; Mismatches 17; Indels 7; Gaps 2;

Qy 1 MASAOVPSFPLAHLFIHFVFTSTIIHLQQRIVKLOPLSEKELPMTTOMSSGNTESPENR 60
Db 1 MASAOVPSFPLAHLFIHFVFTSTIIHLQQRIVKLOPLSEKEL-QAVQMSSPNAARTDMQ 59
Qy 61 RDSQHGNGELRGMTINSIGRLGNMGGEYATLFLALARMNGRLAFIPASMHNALAPIERI 120
Db 60 Q-----SAKLOGIFTINSIGRLGNMGGEYATLFLALARMNGRLAFIPESMHNALAPIERI 113

```





```
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE cdp-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase (EC
DE 2.4.1.69) (Fragment).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NJL/MSF;
RA Liu Y., Kitano T., Koide T., Shirosaki T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB039122; BAE68646.1; -.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1 1
FT NON_TER 321 321
FT SEQUENCE 321 AA; 36464 MW; FF5304CD150F774A CRC64;

Query Match 75.4%; Score 1532.5; DB 11; Length 321;
Best Local Similarity 88.4%; Pred. No. 4e-131;
Matches 290; Conservative 13; Mismatches 18; Indels 7; Gaps 2;

QY 15 FLIFVFVTSIIHQQRIKVLQPLSEKELPMTQMSSGNTPESPMRRDSQHGNGELRGM 74
DB 1 FLIFVFVTSIIHQQRIKVLQPLSEKELPMTQMSSGNTPESPMRRDSQHGNGELRGM 74
QY 75 FTINSIGRLGNQGEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIP 134
DB 54 FTINSIGRLGNQGEYATLALARMNGRLAFIPESMINALAPIFRISLPVLHSDTARRIP 113
QY 135 WQNYHLNDMMEERYRHIPGHEVFTGYPCSWTFYHHLRPEILKEFTLHDHVREREAQAF 194
DB 114 WQNYHLNDMMEERYRHIPGHEVFTGYPCSWTFYHHLRPEILKEFTLHDHVREREAQAF 173
QY 195 GLRVNGSQSTFVGVHVVRGDDYVHVMPKVGVDVADRGYLEKALDMFRARYSSPVFVVT 254
DB 174 GLRVNGSQSTFVGVHVVRGDDYVHVMPKVGVDVADRGYLEKALDMFRARYSSPVFVVT 233
QY 255 NGMAWCRENINASRGDVFVFNAGNIEGSPAKDFALLTCNHTIMTIGTFGWAAYLAGDGT 314
DB 234 NGMAWCRENINISUGDVFVFNAGNIEGSPAKDFALLTCNHTIMTIGTFGWAAYLAGDGT 293
QY 315 IYLANITLPSDPLKVPKPAALPEWV 342
DB 294 IYLANITLPSDPLKVPKPAALPEWV 321

RESULT 10
O77486 PRELIMINARY; PRT; 343 AA.
AC O77486;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Alpha(1,2) fucosyltransferase.
GN FUT2.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;
RT "The old origin of a null allele se428 of the human ABO-secretor type
RT alpha(1,2) fucosyltransferase gene (FUT2).";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

Query Match 73.6%; Score 1495.5; DB 6; Length 343;
```

```
RX MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human secretor blood
RT group alpha(1,2) fucosyltransferase gene (FUT2).";
RL J. Biol. Chem. 270:4640-4649(1995).
DR EMBL: AB015635; BAA31128.1; -.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 343 AA; 38973 MW; 894E28BD74AE8FBC CRC64;

Query Match 73.8%; Score 1499.5; DB 6; Length 343;
Best Local Similarity 78.5%; Pred. No. 4.4e-128;
Matches 277; Conservative 29; Mismatches 36; Indels 11; Gaps 2;

QY 1 MASQVDFSPPLAHLFLIPFVFTSTIIHQQRIKVLQPLSEKELPMTQMSSGNTPESPMR 60
DB 1 MLVYQMPDFSPMAHFLFVFTSTIFHVQQRLAKIQAM--WELEFV-----QIEPLA 49
QY 61 RDSBOHNGELRGMTFINSIGRLGNQGEYATLALARMNGRLAFIPASMHNALAPIFR 120
DB 50 STSKALGPSQLRGWTTINAGRLGNQGEYATLALAKMNGRPAPFPAQMHTLAPIFR 109
QY 121 SLPVLHSDTAKKIPWQNYHLNDMMEERYRHIPGHEVFTGYPCSWTFYHHLRPEILKEFT 180
DB 110 TLPVLHSATASRIPWQNYHLNDMMEERYRHIPGHEVFTGYPCSWTFYHHLRQELQFT 169
QY 181 LHDHVREREAQAFGLRVNGSQSTFVGVHVVRGDDYVHVMPKVGVDVADRGYLEKALDM 240
DB 170 LHDHVREREAQAFGLRVNGSQSTFVGVHVVRGDDYVHVMPKVGVDVADRYLQQLDW 229
QY 241 FRARYSSPVFVVTNGMAWCRENINASRGDVFVFNAGNIEGSPAKDFALLTCNHTIMTIG 300
DB 230 FRARYSSPIFVVTNGMAWCRENIDTSHGDVVFVFNAGNIEGSPAKDFALLTCNHTIMTIG 289
QY 301 TFGWAAYLAGDGTIYLANITLPSDPLKVPKPAALPEWVGIPADLSPLK 353
DB 290 TFGWAAYLAGDGTIYLANITLPSDPLKVPKPAALPEWVGIPADLSPLK 342

RESULT 11
O77485 PRELIMINARY; PRT; 343 AA.
AC O77485;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Alpha(1,2) fucosyltransferase.
GN FUT2.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;
RT "The old origin of a null allele se428 of the human ABO-secretor type
RT alpha(1,2) fucosyltransferase gene (FUT2).";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human secretor blood
RT group alpha(1,2) fucosyltransferase gene (FUT2).";
RL J. Biol. Chem. 270:4640-4649(1995).
DR EMBL: AB015634; BAA31127.1; -.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 343 AA; 38999 MW; B849D7E152852081 CRC64;

Query Match 73.6%; Score 1495.5; DB 6; Length 343;
```

```

Best Local Similarity 78.2%; Pred. No. 1e-127;
Matches 276; Conservative 31; Mismatches 35; Indels 11; Gaps 2;

QY 1 MASQVPPSPFLAHLIFVVTSTIIHLQORIVKLOPLSEKELPMTTQMSSGNTSEPMR 60
Db 1 MLVVQMPFSPFAHFLIFVVTSTIIHVQORLAKIQAM--WELPV-----QIPVLA 49

QY 61 RDSQHGNGELRGMTINSIGRLGNQMGEXATLAFALARMNGRLAFIPASMHNALAFIRI 120
Db 50 STSKALGPSQLRGMTINAGRLGNQMGEXATLYALAKMNGRPAFIPAQMSTLAFIRI 109

QY 121 SLPLVHSDTAKKIPWQNYHLNDWMEERYRHIPGHVFRTGYPGCSWTFFYHHLRPEILKEFT 180
Db 110 TLPVLHSATASRIPWQNYHLNDWMEERYRHIPGCVYRTGYPGCSWTFFYHHLRQEILOEFT 169

QY 181 LHDHVRREAQAFRLGLRVNGSQPSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDM 240
Db 170 LHDHVRREAQAFRLGLQVNGSRPGTFVGHVRRGDYVHVMPKVKWGVVADRRYLOQALDW 229

QY 241 FRARYSSPFVVTSGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTCOCHNTIMTIG 300
Db 230 FRARYSSPIFVVTSGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTCOCHNTIMTIG 289

QY 301 TFGTAAAYLAGGDTIYLANTYLDPSPKLVKFKPEAAFLPEWVGIPADLSPLK 353
Db 290 TFGTAAAYLTGGDTIYLANTYLDPSPKLVKFKPEAAFLPEWVGIAADLSPLK 342

RESULT 12
Q9TUD3 PRELIMINARY; PRT; 343 AA.
AC Q9TUD3;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE Alpha (1.2) fucosyl transferase.
GN FUT2
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ALEXIS;
RX MEDLINE=20188794; PubMed=10723735;
RA Apoll P.-A., Roubinet F., Desplau S., Mollicone R., Oriol R.,
RA Blancher A.;
RT "Evolution of alpha2-Fucosyltransferase genes in primates: relation
RT between an intronic Alu-Y element and red cell expression of ABH
RT antigens.";
RL Mol. Biol. Evol. 17:337-351(2000).
DR EMBL; AF080606; AAF14068.1;
DR InterPro; IPR002516; GT.11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Transferase.
SQ SEQUENCE 343 AA; 39001 MW; A753375D47AE8C6C CRC64;

Query Match 73.6%; Score 1495.5; DB 6; Length 343;
Best Local Similarity 78.2%; Pred. No. 1e-127;
Matches 276; Conservative 30; Mismatches 36; Indels 11; Gaps 2;

QY 1 MASQVPPSPFLAHLIFVVTSTIIHLQORIVKLOPLSEKELPMTTQMSSGNTSEPMR 60
Db 1 MLVVQMPFSPFAHFLIFVVTSTIIHVQORLAKIQAM--WELPV-----QIPVLA 49

QY 61 RDSQHGNGELRGMTINSIGRLGNQMGEXATLAFALARMNGRLAFIPASMHNALAFIRI 120
Db 50 STSKALGPSQLRGMTINAGRLGNQMGEXATLYALAKMNGRPAFIPAQMSTLAFIRI 109

QY 121 SLPLVHSDTAKKIPWQNYHLNDWMEERYRHIPGHVFRTGYPGCSWTFFYHHLRPEILKEFT 180
Db 110 TLPVLHSATASRIPWQNYHLNDWMEERYRHIPGCVYRTGYPGCSWTFFYHHLRQEILOEFT 169

```

```

QY 181 LHDHVRREAQAFRLGLRVNGSQPSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDM 240
Db 170 LHDHVRREAQAFRLGLQVNGSRPGTFVGHVRRGDYVHVMPKVKWGVVADRRYLOQALDW 229

QY 241 FRARYSSPFVVTSGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTCOCHNTIMTIG 300
Db 230 FRARYSSPIFVVTSGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTCOCHNTIMTIG 289

QY 301 TFGTAAAYLAGGDTIYLANTYLDPSPKLVKFKPEAAFLPEWVGIPADLSPLK 353
Db 290 TFGTAAAYLTGGDTIYLANTYLDPSPKLVKFKPEAAFLPEWVGIAADLSPLK 342

RESULT 13
Q9TUD5 PRELIMINARY; PRT; 338 AA.
AC Q9TUD5;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE Alpha (1.2) fucosyl transferase.
GN FUT2
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHI6;
RX MEDLINE=20188794; PubMed=10723735;
RA Apoll P.-A., Roubinet F., Desplau S., Mollicone R., Oriol R.,
RA Blancher A.;
RT "Evolution of alpha2-Fucosyltransferase genes in primates: relation
RT between an intronic Alu-Y element and red cell expression of ABH
RT antigens.";
RL Mol. Biol. Evol. 17:337-351(2000).
DR EMBL; AF080604; AAF14066.1;
DR InterPro; IPR002516; GT.11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Transferase.
SQ SEQUENCE 338 AA; 38428 MW; E7FABA0FF1BC95F9 CRC64;

Query Match 73.3%; Score 1488.5; DB 6; Length 338;
Best Local Similarity 78.7%; Pred. No. 4.3e-127;
Matches 274; Conservative 31; Mismatches 32; Indels 11; Gaps 2;

QY 6 VPSFPLAHLIFVVTSTIIHLQORIVKLOPLSEKELPMTTQMSSGNTSEPMRDSQ 65
Db 1 MPFSPPVAHFLIFVVTSTIIHVQORLAKIQAM--WELPV-----QIPVLA 49

QY 66 HGNGELRGMTINSIGRLGNQMGEXATLAFALARMNGRLAFIPASMHNALAFIRISLPVL 125
Db 50 LGPSQLRGMTINAGRLGNQMGEXATLYALAKMNGRPAFIPAQMSTLAFIRISLPVL 109

QY 126 HSDTAKKIPWQNYHLNDWMEERYRHIPGHVFRTGYPGCSWTFFYHHLRPEILKEFTLHGHV 185
Db 110 HSATASRIPWQNYHLNDWMEERYRHIPGCVYRTGYPGCSWTFFYHHLRQEILOEFTLHGHV 169

QY 186 REBAQAFRLGLRVNGSQPSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDMFEARY 245
Db 170 REBAQAFRLGLQVNGSRPGTFVGHVRRGDYVHVMPKVKWGVVADRRYLOQALDWPRARY 229

QY 246 SSPVVTSGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTCOCHNTIMTIGTIGIW 305
Db 230 SSPVVTSGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTCOCHNTIMTIGTIGIW 289

QY 306 AAYLAGGDTIYLANTYLDPSPKLVKFKPEAAFLPEWVGIPADLSPLK 353
Db 290 AAYLTGGDTIYLANTYLDPSPKLVKFKPEAAFLPEWVGIAADLSPLK 337

RESULT 14
Q9T505

```



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:01:52 : Search time 38.17 seconds  
(without alignments)  
1932.533 Million cell updates/sec

Title: US-10-040-863-8\_COPY\_23\_380

Perfect score: 1921

Sequence: 1 STIIHQRIKVLQPLSEKE.....RSHFHLKAGVTCYVAGRAF 358

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_protist.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1921	100.0	380	11	Q9JK44
2	1896	98.7	353	11	Q55025
3	1772	92.2	354	11	Q35087
4	1767	92.0	354	11	Q9R275
5	1563.5	81.4	347	11	Q9JL27
6	1559.5	81.2	347	11	Q70504
7	1508.5	78.5	321	11	Q920W2
8	1508.5	78.5	321	11	Q91VF0
9	1493.5	77.7	321	11	Q920W3
10	1421.5	74.0	343	6	Q77486
11	1418.5	73.8	338	6	Q9TUD5
12	1418.5	73.8	343	6	Q77485
13	1417.5	73.8	343	6	Q9TUD3
14	1412.5	73.5	347	6	Q29505
15	1410.5	73.4	343	6	Q77487
16	1403.5	73.1	332	4	Q99450

#### ALIGNMENTS

RESULT 1

Q9JK44 PRELIMINARY; PRT; 380 AA.  
AC Q9JK44;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Alpha 1-2 fucosyltransferase.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-FISCHER;  
RA Sherwood A.L.; Stroud M.R.; Levery S.B.; Holmes E.H.;  
RT "An amino acid region at the N-terminus of rat hepatoma H35 cell GDP-fucose:GMI alpha 1-2 fucosyltransferase is required for optimum enzyme activity and interaction with lipids."  
RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF264005; AAF72200.1;  
DR InterPro; IPR002516; GT\_11;  
DR Pfam; PF01531; Glyco\_transf\_11; 1.  
DR Glycosyltransferase; Transferase.  
SQ SEQUENCE 380 AA; 42742 MW; 920692C96689C05C CRC64;

Query Match 100.0%; Score 1921; DB 11; Length 380;  
Best Local Similarity 100.0%; Pred. No. 1.2e+166;  
Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIIHQRIKVLQPLSEKELPMTTQSSGNTESPMDRDSQHGNGELRGMTINSIGR 60  
DB 23 STIIHQRIKVLQPLSEKELPMTTQSSGNTESPMDRDSQHGNGELRGMTINSIGR 82  
QY 61 LGNQMGEYATLFAALRMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQWYHLND 120  
DB 83 LGNQMGEYATLFAALRMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQWYHLND 142  
QY 121 KWEERYHPTGHEVRFYGYCSWTFYHHLRPEILKEFTLHDHVREAAQAFILRLGRVNGSQ 180

Db 143 WMEERYRHIPGHEVFTGTCPSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNSQ 202  
 QY 161 PSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240  
 Db 203 PSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 262  
 QY 241 NINASRGDGVFAGNIEGSPAKDFALLTQCNTHTMTIGTFGIWAAYLAGGDTIYLYANTL 300  
 Db 263 NINASRGDGVFAGNIEGSPAKDFALLTQCNTHTMTIGTFGIWAAYLAGGDTIYLYANTL 322  
 QY 301 PDSPLKVKFPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 358  
 Db 323 PDSPLKVKFPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 380

RESULT 2  
 O55025 PRELIMINARY: PRT: 353 AA.  
 AC O55025;  
 DT 01-JUN-1998 (Tremblrel. 06, Created)  
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Alpha 1,2 fucosyltransferase (fragment).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BUFFALO RAT;  
 RX MEDLINE=98342056; PubMed=9675030;  
 RA "Shewood A.L., Holmes B.H.;  
 RT Cloning and expression of the catalytic domain from rat hepatoma H35  
 RT cell Gdp-fucose: GM1 alpha 1->2fucosyltransferase, an enzyme which is  
 RT activated during early stages of chemical carcinogenesis in rat  
 RT liver.";  
 RL Arch. Biochem. Biophys. 355:215-221(1998).  
 DR EMBL; AF042743; AAC14695.1; -;  
 DR InterPro; IPR002516; GT\_11.  
 DR Pfam; PF01531; Glyco.transf.11; 1.  
 KW Glycosyltransferase; Transferase.  
 FT NON\_TER  
 SQ SEQUENCE 353 AA; 39738 MW; 609D64EB222C9585 CRC64;

Query Match 98.7%; Score 1896; DB 11; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-164;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 LOORIVKLOPSEKELPMTQSSGNTSEPMRRDSEHGNGELRGMTINSIGRLGNM 65  
 Db 1 LOORIVKLOPSEKELPMTQSSGNTSEPMRRDSEHGNGELRGMTINSIGRLGNM 60  
 QY 66 GEYATLPALARMNGRLAIPASMHNALAPIRISLPVLHSDTAKKIPQNYHLNDWMEER 125  
 Db 61 GEYATLPALARMNGRLAIPASMHNALAPIRISLPVLHSDTAKKIPQNYHLNDWMEER 120  
 QY 126 YRHIPGHEVFTGTCPSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNSQSPSTFV 185  
 Db 121 YRHIPGHEVFTGTCPSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNSQSPSTFV 180  
 QY 186 GVHVRRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCREINAS 245  
 Db 181 GVHVRRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCREINAS 240  
 QY 246 RGVVVFAGNIEGSPAKDFALLTQCNTHTMTIGTFGIWAAYLAGGDTIYLYANTLPDPSFF 305  
 Db 241 RGVVVFAGNIEGSPAKDFALLTQCNTHTMTIGTFGIWAAYLAGGDTIYLYANTLPDPSFF 300  
 QY 306 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 358  
 Db 301 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 353

RESULT 3  
 O35087 PRELIMINARY: PRT: 354 AA.  
 AC O35087;  
 DT 01-JAN-1998 (Tremblrel. 05, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Alpha 1,2-fucosyltransferase.  
 GN FTB.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=COLON CANCER;  
 RA Soejima M., Wang B., Koda Y., Kimura H.;  
 RT "Two distinct rat Gdp-L-fucose:b-D-galactoside 2-a-L-  
 RT fucosyltransferase genes";  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB006138; BAA21742.1; -;  
 DR InterPro; IPR002516; GT\_11.  
 DR Pfam; PF01531; Glyco.transf.11; 1.  
 KW Glycosyltransferase; Transferase.  
 SQ SEQUENCE 354 AA; 39995 MW; 8636444488215BA1 CRC64;  
 Query Match 92.2%; Score 1772; DB 11; Length 354;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-153;  
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 STIIHLQQRIVKLOPSEKELPMTQSSGNTSEPMRRDSEHGNGELRGMTINSIGR 60  
 Db 23 STIIHLQQRIVKLOPSEKELPMTQSSGNTSEPMRRDSEHGNGELRGMTINSIGR 82  
 QY 61 LGNQMEYATLPALARMNGRLAIPASMHNALAPIRISLPVLHSDTAKKIPQNYHLND 120  
 Db 83 LGNQMEYATLPALARMNGRLAIPASMHNALAPIRISLPVLHSDTAKKIPQNYHLND 142  
 QY 121 WMEERYRHIPGHEVFTGTCPSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNSQ 180  
 Db 143 WMEERYRHIPGHEVFTGTCPSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNSQ 202  
 QY 181 PSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240  
 Db 203 PSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 262  
 QY 241 NINASRGDGVFAGNIEGSPAKDFALLTQCNTHTMTIGTFGIWAAYLAGGDTIYLYANTL 300  
 Db 263 NINASRGDGVFAGNIEGSPAKDFALLTQCNTHTMTIGTFGIWAAYLAGGDTIYLYANTL 322  
 QY 301 PDSPLKVKFPEAAFLPEWVGIPADLSPLLK 331  
 Db 323 PDSPLKVKFPEAAFLPEWVGIPADLSPLLK 353  
 RESULT 4  
 O9R275 PRELIMINARY: PRT: 354 AA.  
 AC O9R275;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Alpha 1,2-fucosyltransferase B.  
 GN FTB.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BDIX;  
 RX MEDLINE=21099374; PubMed=11179967;

```

RA Bureau V., Marionneau S., Cailleau-Thomas A., Le Moullac-Vaidre B.,
RA Liehr T., Le Pendu J.;
RT "Comparison of the three rat GDP-L-fucose: beta-D-galactoside 2-
RT alpha-L-fucosyltransferases FTA, FTB and FTC.";
RL Eur. J. Biochem. 268:1006-1019(2001).
DR EMBL: AF131238; AAD24469.1; -.
DR InterPro: IPR002516; GT_11;
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 354 AA; 39983 MW; 123B8C8379B8559E CRC64;

Query Match 92.0%; Score 1767; DB 11; Length 354;
Best Local Similarity 99.7%; Pred. No. 1.2e-152;
Matches 330; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STIIHQORIVKLOPLSEKELPMTTQMSSNGNTSEPMRRDSEQHNGELRGMTINSIGR 60
Db 23 STIIHQORIVKLOPLSEKELPMTTQMSSNGNTSEPMRRDSEQHNGELRGMTINSIGR 75
QY 61 LGNOMGEYATLAFALARMNGRLAFIPASMHNALAFIFRISLPVLHSDTAKKIPWQNYHLND 120
Db 83 LGNOMGEYATLAFALARMNGRLAFIPASMHNALAFIFRISLPVLHSDTAKKIPWQNYHLND 142
QY 121 WMEERYRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQ 180
Db 143 WMEERYRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQ 202
QY 181 PSTFVGVRGDDVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240
Db 203 PSTFVGVRGDDVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 262
QY 241 NINASRGDGVFAGNGIEGSPAKDFALLTOCNHTTMTGTGFIWAAYLAGGDTIYLANVTL 300
Db 263 NINASRGDGVFAGNGIEGSPAKDFALLTOCNHTTMTGTGFIWAAYLAGGDTIYLANVTL 322
QY 301 PDSFPLKVFKEPAEAFLEPWGIPADLSPLK 331
Db 323 PDSFPLKVFKEPAEAFLEPWGIPADLSPLK 353

RESULT 5
Q9JL27 PRELIMINARY; PRT; 347 AA.
AC Q9JL27;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Alpha(1,2)fucosyltransferase FUT2 (EC 2.4.1.69).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/OLA;
RX MEDLINE=21316545; PubMed=11323419;
RA Domino S.E., Zhang L., Lowe J.B.;
RT "Molecular Cloning, Genomic Mapping, and Expression of Two Secretor
RT Blood Group alpha (1,2)fucosyltransferase Genes Differentially
RT Regulated in Mouse Uterine Epithelium and Gastrointestinal Tract.";
RL J. Biol. Chem. 276:23748-23756(2001).
DR EMBL: AF214656; AAF45146.1; -.
DR MGD: MGI:109374; Fut2.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 347 AA; 39243 MW; DD5F441046CC8E79 CRC64;

Query Match 81.4%; Score 1563.5; DB 11; Length 347;
Best Local Similarity 89.1%; Pred. No. 4.1e-134;
Matches 295; Conservative 12; Mismatches 17; Indels 7; Gaps 2;

```

```

QY 1 STIIHQORIVKLOPLSEKELPMTTQMSSNGNTSEPMRRDSEQHNGELRGMTINSIGR 60
Db 23 STIIHQORIVKLOPLSEKELPMTTQMSSNGNTSEPMRRDSEQHNGELRGMTINSIGR 75
QY 61 LGNOMGEYATLAFALARMNGRLAFIPASMHNALAFIFRISLPVLHSDTAKKIPWQNYHLND 120
Db 76 LGNOMGEYATLAFALARMNGRLAFIPESMHNALAFIFRISLPVLHSDTARRIPWQNYHLND 135
QY 121 WMEERYRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQ 180
Db 136 WMEERYRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQ 195
QY 181 PSTFVGVRGDDVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240
Db 196 PSTFVGVRGDDVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 255
QY 241 NINASRGDGVFAGNGIEGSPAKDFALLTOCNHTTMTGTGFIWAAYLAGGDTIYLANVTL 300
Db 256 NINASRGDGVFAGNGIEGSPAKDFALLTOCNHTTMTGTGFIWAAYLAGGDTIYLANVTL 315
QY 301 PDSFPLKVFKEPAEAFLEPWGIPADLSPLK 331
Db 316 PDSFPLKVFKEPAEAFLEPWGIPADLSPLK 346

RESULT 6
Q70504 PRELIMINARY; PRT; 347 AA.
AC Q70504;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase (EC
DE 2.4.1.69).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=GASTROINTESTINAL TRACT;
RX MEDLINE=20471982; PubMed=11018479;
RA Lin B., Hayashi Y., Saito M., Sakakibara Y., Yanagisawa M.,
RA Ikamori M.;
RT "GDP-fucose: beta-galactoside alpha1,2-fucosyltransferase, MFUT-II,
RT and not MFUT-I or -III, is induced in a restricted region of the
RT digestive tract of germ-free mice by host-microbe interactions and
RT cycloheximide.";
RL Biochim. Biophys. Acta 1487:275-285(2000).
DR EMBL: AF064792; AAC16887.1; -.
DR MGD: MGI:109374; Fut2.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 347 AA; 39215 MW; 616CCB11581C4179 CRC64;

Query Match 81.2%; Score 1559.5; DB 11; Length 347;
Best Local Similarity 88.8%; Pred. No. 9.4e-134;
Matches 294; Conservative 13; Mismatches 17; Indels 7; Gaps 2;

QY 1 STIIHQORIVKLOPLSEKELPMTTQMSSNGNTSEPMRRDSEQHNGELRGMTINSIGR 60
Db 23 STIIHQORIVKLOPLSEKELPMTTQMSSNGNTSEPMRRDSEQHNGELRGMTINSIGR 75
QY 61 LGNOMGEYATLAFALARMNGRLAFIPASMHNALAFIFRISLPVLHSDTAKKIPWQNYHLND 120
Db 76 LGNOMGEYATLAFALARMNGRLAFIPESMHNALAFIFRISLPVLHSDTARRIPWQNYHLND 135
QY 121 WMEERYRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQ 180
Db 136 WMEERYRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQ 195

```

```

QY 181 PSTFVGVHVRGDIYVHVPNVKGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCRE 240
Db 196 PSTFVGVHVRGDIYVHVPNVKGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCRE 255
QY 241 NNASRGDVVFAGNGIEGSPAKDFALLTQCNTTMTIGTFGIWAAYLAGGDIYIANYTL 300
Db 256 NNTSLGDDVVFAGNGIEGSPAKDFALLTQCNTTMTIGTFGIWAAYLAGGDIYIANYTL 315
QY 301 PDSPLKVKFKPAAPLPEWVGPADLSPLLK 331
Db 316 PDSPLKVKFKPAAPLPEWVGPADLSPLLK 346
RESULT 7
Q920W2
ID Q920W2 PRELIMINARY; PRT; 321 AA.
AC Q920W2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase (EC 2.4.1.69) (Fragment).
DE 2.4.1.69) (Fragment).
GN FUT2.
OS Mus spicilegus (Stepe mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZBN;
RA Liu Y., Kitano T., Koide T., Shiroishi T., Moriwaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039123; BAB68647.1; -
DR EMBL; AB039123; BAB68647.1; -
DR InterPro: IPR002516; GT_11
DR Pfam: PF01531; Glyco_Transf_11; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1
FT NON_TER 321
SQ SEQUENCE 321 AA; 36560 MW; 5B7D19BA6D4B5394 CRC64;
Query Match 78.5%; Score 1508.5; DB 11; Length 321;
Best Local Similarity 89.1%; Pred. No. 3.7e-129;
Matches 285; Conservative 11; Mismatches 17; Indels 7; Gaps 2;
QY 1 STIIHQQRIVKVLQPLSEKELPMTTQMSSGNTSPESPMRRDSEQHNGELRGMTINSIGR 60
Db 9 STIIHQQRIVKVLQPLSEKEL-QAQMSSLNARTEMQQ-----SAKLOGIFTINSIGR 61
QY 61 LGNOMGEYATLALARMNGRLAFIPASHMNALAPIFRISLPVLHSDTAKKIPQWNYHLND 120
Db 62 LGNOMGEYATLALARMNGRLAFIPASHMNALAPIFRISLPVLHSDTARRIPQWNYHLND 121
QY 121 WHEERYRHIPGHEVFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQ 180
Db 122 WHEERYRHIPGHEVFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQ 181
QY 181 PSTFVGVHVRGDIYVHVPNVKGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCRE 240
Db 182 PSTFVGVHVRGDIYVHVPNVKGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCRE 241
QY 241 NNASRGDVVFAGNGIEGSPAKDFALLTQCNTTMTIGTFGIWAAYLAGGDIYIANYTL 300
Db 242 NNTSLGDDVVFAGNGIEGSPAKDFALLTQCNTTMTIGTFGIWAAYLAGGDIYIANYTL 301
QY 301 PDSPLKVKFKPAAPLPEWV 320
Db 302 PDSPLKVKFKPAAPLPEWV 321
RESULT 8
Q91VFO
ID Q91VFO PRELIMINARY; PRT; 321 AA.
AC Q91VFO;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase (EC 2.4.1.69) (Fragment).
DE 2.4.1.69) (Fragment).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VARIOUS STRAINS;
RA Liu Y., Kitano T., Koide T., Shiroishi T., Moriwaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039114; BAB68638.1; -
DR EMBL; AB039114; BAB68638.1; -
DR EMBL; AB039115; BAB68639.1; -
DR EMBL; AB039115; BAB68640.1; -
DR EMBL; AB039117; BAB68641.1; -
DR EMBL; AB039118; BAB68642.1; -
DR EMBL; AB039119; BAB68643.1; -
DR EMBL; AB039120; BAB68644.1; -
DR EMBL; AB039121; BAB68645.1; -
DR InterPro: IPR002516; GT_11
DR Pfam: PF01531; Glyco_Transf_11; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1
FT NON_TER 321
SQ SEQUENCE 321 AA; 36516 MW; 0622D3B503B72D1 CRC64;
Query Match 78.5%; Score 1508.5; DB 11; Length 321;
Best Local Similarity 88.8%; Pred. No. 3.7e-129;
Matches 284; Conservative 12; Mismatches 17; Indels 7; Gaps 2;
QY 1 STIIHQQRIVKVLQPLSEKELPMTTQMSSGNTSPESPMRRDSEQHNGELRGMTINSIGR 60
Db 9 STIIHQQRIVKVLQPLSEKEL-QAQMSSLNARTEMQQ-----SAKLOGIFTINSIGR 61
QY 61 LGNOMGEYATLALARMNGRLAFIPASHMNALAPIFRISLPVLHSDTAKKIPQWNYHLND 120
Db 62 LGNOMGEYATLALARMNGRLAFIPASHMNALAPIFRISLPVLHSDTARRIPQWNYHLND 121
QY 121 WHEERYRHIPGHEVFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQ 180
Db 122 WHEERYRHIPGHEVFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQ 181
QY 181 PSTFVGVHVRGDIYVHVPNVKGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCRE 240
Db 182 PSTFVGVHVRGDIYVHVPNVKGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCRE 241
QY 241 NNASRGDVVFAGNGIEGSPAKDFALLTQCNTTMTIGTFGIWAAYLAGGDIYIANYTL 300
Db 242 NNTSLGDDVVFAGNGIEGSPAKDFALLTQCNTTMTIGTFGIWAAYLAGGDIYIANYTL 301
QY 301 PDSPLKVKFKPAAPLPEWV 320
Db 302 PDSPLKVKFKPAAPLPEWV 321
RESULT 9
Q920W3
ID Q920W3 PRELIMINARY; PRT; 321 AA.
AC Q920W3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase (EC 2.4.1.69) (Fragment).
DE 2.4.1.69) (Fragment).
GN FUT2.

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NUL/MSF;
RA Liu Y., Kitano T., Koide T., Shirolshi T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039122; BAB68646.1; -.
DR InterPro: IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1
FT NON_TER 321
SQ SEQUENCE 321 AA; 36464 MW; FF5304CD150F774A CRC64;

Query Match 77.7%; Score 1493.5; DB 11; Length 321;
Best Local Similarity 88.1%; Pred. No. 8.7e-128;
Matches 282; Conservative 13; Mismatches 18; Indels 7; Gaps 2;

QY 1 STIHLQORIVKLPSEKELPMTQSSGNTESEPMRDSSEHQNGELRGMTINSIGR 60
Db 9 STIHLQORIVKLPSEKELPMTQSSGNTESEPMRDSSEHQNGELRGMTINSIGR 61
QY 61 LGNQMGYATLAFALARNGLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLND 120
Db 62 LGNQMGYATLAFALARNGLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLND 121
QY 121 WMEERYRHIPGHFVRFTGYPCSWTFYHHLRPEILKERTLHDHVREEAQAFRLGLRVNGSQ 180
Db 122 WMEERYRHIPGHFVRFTGYPCSWTFYHHLRPEILKERTLHDHVREEAQAFRLGLRVNGSQ 181
QY 181 PSTFVGHVHRRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240
Db 182 PSTFVGHVHRRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 241
QY 241 NINASRGDGVFAGNGIEGSPAKDFALLTQCNHTTMTGTGFINAAAYLAGDITIYLANITL 300
Db 242 NINASRGDGVFAGNGIEGSPAKDFALLTQCNHTTMTGTGFINAAAYLAGDITIYLANITL 301
QY 301 PDSPLKLVFKPEAAFLPEW 320
Db 302 PDSPLKLVFKPEAAFLPEW 321

RESULT 10
O77486 PRELIMINARY; PRT; 343 AA.
AC O77486;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha(1,2) fucosyltransferase.
GN FUT2.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;
RT "The old origin of a null allele se428 of the human ABO-secretor type
RT alpha(1,2) fucosyltransferase gene (FUT2).";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human secretor blood
RT group alpha(1,2) fucosyltransferase gene (FUT2).";

```

```

J. Biol. Chem. 270:4640-4649(1995).
RL EMBL; AB015635; BAA31128.1; -.
DR InterPro: IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 343 AA; 38973 MW; 894E28BD74AE8FBC CRC64;

Query Match 74.0%; Score 1421.5; DB 6; Length 343;
Best Local Similarity 79.8%; Pred. No. 3.5e-121;
Matches 264; Conservative 25; Mismatches 31; Indels 11; Gaps 2;

QY 1 STIHLQORIVKLPSEKELPMTQSSGNTESEPMRDSSEHQNGELRGMTINSIGR 60
Db 23 STIHLQORIVKLPSEKELPMTQSSGNTESEPMRDSSEHQNGELRGMTINSIGR 61
QY 61 LGNQMGYATLAFALARNGLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLND 120
Db 72 LGNQMGYATLAFALARNGLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLND 131
QY 121 WMEERYRHIPGHFVRFTGYPCSWTFYHHLRPEILKERTLHDHVREEAQAFRLGLRVNGSQ 180
Db 132 WMEERYRHIPGHFVRFTGYPCSWTFYHHLRPEILKERTLHDHVREEAQAFRLGLRVNGSQ 191
QY 181 PSTFVGHVHRRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240
Db 192 PSTFVGHVHRRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 251
QY 241 NINASRGDGVFAGNGIEGSPAKDFALLTQCNHTTMTGTGFINAAAYLAGDITIYLANITL 300
Db 252 NINASRGDGVFAGNGIEGSPAKDFALLTQCNHTTMTGTGFINAAAYLAGDITIYLANITL 311
QY 301 PDSPLKLVFKPEAAFLPEW 331
Db 312 PDSPLKLVFKPEAAFLPEW 342

RESULT 11
Q9TUD5 PRELIMINARY; PRT; 338 AA.
AC Q9TUD5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha(1,2) fucosyltransferase.
GN FUT2.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CH16.
RX MEDLINE=20188794; PubMed=10723735;
RA Apoll P.-A., Roubinet F., Despiau S., Mollicone R., Oriol R.,
RA Blancher A.;
RT "Evolution of alpha2-Fucosyltransferase genes in primates: relation
RT between an intronic Alu-Y element and red cell expression of ABH
RT antigens.";
RL Mol. Biol. Evol. 17:337-351(2000).
DR EMBL; AF080604; AAF14066.1; -.
DR InterPro: IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Transferase
SQ SEQUENCE 338 AA; 36428 MW; E7FADA0FF1BC95F9 CRC64;

Query Match 73.8%; Score 1418.5; DB 6; Length 338;
Best Local Similarity 79.5%; Pred. No. 6.4e-121;
Matches 263; Conservative 27; Mismatches 30; Indels 11; Gaps 2;

QY 1 STIHLQORIVKLPSEKELPMTQSSGNTESEPMRDSSEHQNGELRGMTINSIGR 60
Db 18 STIHLQORIVKLPSEKELPMTQSSGNTESEPMRDSSEHQNGELRGMTINSIGR 66

```



```

QY 61 LGNQGEYATLALARMNGRLAFIPASMHNALAPIRISLPVLSHDTAKKIPWQNYHLND 120
Db 67 LGNQGEYATLALAKMNGRPAPFAQMHSTLAPFRITLPVLSHATASRIPWQNYHLND 126
QY 121 WMEERYRHIPGHVFRFTGYPCSWTYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQ 180
Db 127 WMEERYRHIPGHVFRFTGYPCSWTYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSR 186
QY 181 PSTFVGHVRRGDYVHVMPKWKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCREE 240
Db 187 PSTFVGHVRRGDYVHVMPKWKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCREE 246
QY 241 NINASRGDVVFAGNGIEGSPAKDFALLTCQCNHTIMTIGTFGIIWAAYLAGGDIYILANYTL 300
Db 247 NIDTSHGDVVFAGDGIEGSPAKDFALLTCQCNHTIMTIGTFGIIWAAYLAGGDIYILANYTL 306
QY 301 PDSPLKVKFPEAAFLPEWVGIPADLSPLK 331
Db 307 PDSPLKVKFPEAAFLPEWVGIAADLSPLK 337

RESULT 12
Q77485
ID O77485 PRELIMINARY: PRT: 343 AA.
AC O77485;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha(1,2)fucosyltransferase.
GN FUT2.
OS Eukaryotes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;
RT "The old origin of a null allele se428 of the human ABO-secretor type
RT alpha(1,2) fucosyltransferase gene (FUT2).";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human secretor blood
RT group alpha(1,2) fucosyltransferase gene (FUT2).";
RL J. Biol. Chem. 270:4640-4649(1995).
RL EMBL: AB015634; BAA31127.1; -.
DR InterPro: IPR002516; GI_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 343 AA; 38999 MW; B849D7E152852081 CRC64;

```

```

Query Match 73.8%; Score 1418.5; DB 6; Length 343;
Best Local Similarity 79.5%; Pred. No. 6.5e-121;
Matches 263; Conservative 27; Mismatches 30; Indels 11; Gaps 2;

QY 1 STIIHQQRIVKLQPLSEKELPMTTQMSSGNTSEPMRRDSEQHNGELRGMFTINSIGR 60
Db 23 STIFVQORLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGWWTINAIGR 71
QY 61 LGNQGEYATLALARMNGRLAFIPASMHNALAPIRISLPVLSHDTAKKIPWQNYHLND 120
Db 72 LGNQGEYATLALAKMNGRPAPFAQMHSTLAPFRITLPVLSHATASRIPWQNYHLND 131
QY 121 WMEERYRHIPGHVFRFTGYPCSWTYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQ 180
Db 132 WMEERYRHIPGHVFRFTGYPCSWTYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSR 191
QY 181 PSTFVGHVRRGDYVHVMPKWKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCREE 240
Db 192 PSTFVGHVRRGDYVHVMPKWKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCREE 251

RESULT 14
Q29505
ID Q29505 PRELIMINARY: PRT: 347 AA.
AC Q29505;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)

```

```

QY 241 NINASRGDVVFAGNGIEGSPAKDFALLTCQCNHTIMTIGTFGIIWAAYLAGGDIYILANYTL 300
Db 252 NIDTSHGDVVFAGDGIEGSPAKDFALLTCQCNHTIMTIGTFGIIWAAYLAGGDIYILANYTL 311
QY 301 PDSPLKVKFPEAAFLPEWVGIPADLSPLK 331
Db 312 PDSPLKVKFPEAAFLPEWVGIAADLSPLK 342

RESULT 13
Q9TUD3
ID Q9TUD3 PRELIMINARY: PRT: 343 AA.
AC Q9TUD3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha(1,2)fucosyl transferase.
GN FUT2.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ALEXIS;
RX MEDLINE=20188794; PubMed=10723735;
RA Apoll P.-A., Roubinet F., Despia S., Mollicone R., Oriol R.,
RA Blancher A.;
RT "Evolution of alpha2-Fucosyltransferase genes in primates: relation
RT between an intronic Alu-Y element and red cell expression of ABH
RT antigens.";
RL Mol. Biol. Evol. 17:337-351(2000).
RL EMBL: AF080606; AAF14068.1; -.
DR InterPro: IPR002516; GI_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Transferase.
SQ SEQUENCE 343 AA; 39001 MW; A753375D47AE8C8C CRC64;

Query Match 73.8%; Score 1417.5; DB 6; Length 343;
Best Local Similarity 79.5%; Pred. No. 8.1e-121;
Matches 263; Conservative 26; Mismatches 31; Indels 11; Gaps 2;

QY 1 STIIHQQRIVKLQPLSEKELPMTTQMSSGNTSEPMRRDSEQHNGELRGMFTINSIGR 60
Db 23 STIFVQORLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGWWTINAIGR 71
QY 61 LGNQGEYATLALARMNGRLAFIPASMHNALAPIRISLPVLSHDTAKKIPWQNYHLND 120
Db 72 LGNQGEYATLALAKMNGRPAPFAQMHSTLAPFRITLPVLSHATASRIPWQNYHLND 131
QY 121 WMEERYRHIPGHVFRFTGYPCSWTYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQ 180
Db 132 WMEERYRHIPGHVFRFTGYPCSWTYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSR 191
QY 181 PSTFVGHVRRGDYVHVMPKWKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCREE 240
Db 192 PSTFVGHVRRGDYVHVMPKWKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCREE 251

RESULT 14
Q29505
ID Q29505 PRELIMINARY: PRT: 347 AA.
AC Q29505;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)

```

01-MAR-2002 (TREMELrel. 20, Last annotation update)  
DE Galactoside 2-L-fucosyltransferase 3 (EC 2.4.1.69) (Secretor blood  
DE group ALPHA-2-fucosyltransferase) (GDP-L-fucose:beta-D-galactoside  
DE 2-ALPHA-L-fucosyltransferase) (ALPHA(1,2)F) (Fucosyltransferase 3).  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-GASTROINTESTINAL TRACT;  
RX MEDLINE=96279281; PubMed=8663168;  
RA Hitoshi S., Kojima N., Kanazawa I., Tsuji S.;  
RT "Molecular cloning and expression of a third type of rabbit GDP-L-  
RT fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase.";  
RL J. Biol. Chem. 271:16975-16981(1996)  
CC -!- FUNCTION: CREATES A MEMBRANE-ASSOCIATED PRECURSOR OLIGOSACCHARIDE  
CC FUC-ALPHA((1,2)GALBETA-) CALLED THE H ANTIGEN WHICH IS AN  
CC ESSENTIAL SUBSTRATE FOR THE FINAL STEP IN THE MEMBRANE-ASSOCIATED  
CC A AND B ANTIGEN SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE THE  
CC SAME ACCEPTOR SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.  
CC -!- CATALYTIC ACTIVITY: GDP-L-FUCOSE + BETA-D-GALACTOSYL-R = GDP +  
CC ALPHA-L-FUCOSYL-1,2-BETA-D-GALACTOSYL-R.  
CC -!- PATHWAY: GLYCOSYLATION.  
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND  
CC FORM IN TRANS CISTERNAE OF GOLGI.  
CC -!- TISSUE SPECIFICITY: ADULT GASTROINTESTINAL TRACT AND COLON.  
CC -!- MISCELLANEOUS: THERE ARE THREE GENES (FUT1, FUT2 AND FUT3) WHICH  
CC ENCODE GALACTOSIDE 2-L-FUCOSYLTRANSFERASE IN RABBIT. THEY ARE  
CC EXPRESSED IN A TISSUE-SPECIFIC MANNER.  
CC -!- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN  
CC GLYCOSYLTRANSFERASES.  
DR EMBL: X91269; CAA62669.1; -.  
DR InterPro: IPR002516; GT\_11.  
DR Pfam: PF01531; Glyco\_Transf\_11; 1.  
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;  
KW Signal-anchor; Golgi stack.  
FT DOMAIN 1 5  
FT TRANSMEM 6 28  
FT FT  
FT FT  
FT DOMAIN 29 347  
FT LUMENAL, CATALYTIC (POTENTIAL).  
FT CARBOHYD 192 192  
FT N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 286 286  
FT N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 312 312  
FT N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 347 AA; 39469 MW; 746F7007309862A5 CRC64;

Query Match 73.5%; Score 1412.5; DB 6; Length 347;  
Best Local Similarity 79.8%; Pred. No. 2.3e-120;  
Matches 264; Conservative 25; Mismatches 35; Indels 7; Gaps 1;

QY 1 STIIHQQRIVKLOPLSEKELPMTTQSSNTSPESPMRDSSEOHNGELRGMTINSIGR 60  
DB 23 STIFHQQRIVKLOPLSEKELPMTTQSSNTSPESPMRDSSEOHNGELRGMTINSIGR 75  
QY 61 LGNQMGEXATLFAARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPQWNYHLND 120  
DB 76 LGNQMGEXATLVALAKENGRPAVIPAQMHSTLAFIPRISLPVLHSDTAKKIPQWNYHLND 135  
QY 121 WMEERYRHIPGHEVVRTGPCSWTFYHHLRPEILKFTLHDHVRREAAQALRLGRLVNGSQ 180  
DB 136 WMEERYRHIPGHEVVRTGPCSWTFYHHLRPEILKFTLHDHVRREAAQALRLGRLVNGSR 195  
QY 181 PSTFVGHVHVRGGDYVHVMPPVWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240  
DB 196 PSTFVGHVHVRGGDYVHVMPPVWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 255  
QY 241 NINASRGDGVVFAAGNIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLAGGDTIYLANYYL 300  
DB 256 NIDASRGDGVVFAAGNIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLAGGDTIYLANYYL 315  
QY 301 PDSFFLKVPKPEAAFLPEWVGIPADLSPLIK 331  
DB 316 PDSFFLKVPKPEAAFLPEWVGIPADLSPLIK 346

RESULT 15  
077487 PRELIMINARY; PRT; 343 AA.  
ID O77487  
AC O77487;  
DT 01-NOV-1998 (TREMELrel. 08, Created)  
DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)  
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)  
DE Alpha(1,2) fucosyltransferase.  
GN FUT2.  
OS Pongo pygmaeus (Orangutan).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.  
OX NCBI\_TaxID=9600;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;  
RT "The old origin of a null allele se428 of the human ABO-secretor type  
RT alpha(1,2) fucosyltransferase gene (FUT2).";  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=95181460; PubMed=7876235;  
RX Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;  
RT "Sequence and expression of a candidate for the human secretor blood  
RT group alpha(1,2) fucosyltransferase gene (FUT2).";  
RL J. Biol. Chem. 270:4640-4649(1995).  
DR EMBL: AB015636; BAA31129.1; -.  
DR InterPro: IPR002516; GT\_11.  
DR Pfam: PF01531; Glyco\_Transf\_11; 1.  
KW Glycosyltransferase; Transferase.  
SQ SEQUENCE 343 AA; 38973 MW; 2BD4D2A9704E4A0C CRC64;

Query Match 73.4%; Score 1410.5; DB 6; Length 343;  
Best Local Similarity 79.2%; Pred. No. 3.5e-120;  
Matches 262; Conservative 27; Mismatches 31; Indels 11; Gaps 2;

QY 1 STIIHQQRIVKLOPLSEKELPMTTQSSNTSPESPMRDSSEOHNGELRGMTINSIGR 60  
DB 23 STIFHQQRIVKLOPLSEKELPMTTQSSNTSPESPMRDSSEOHNGELRGMTINSIGR 71  
QY 61 LGNQMGEXATLFAARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPQWNYHLND 120  
DB 72 LGNQMGEXATLVALAKENGRPAVIPAQMHSTLAFIPRISLPVLHSDTAKKIPQWNYHLND 131  
QY 121 WMEERYRHIPGHEVVRTGPCSWTFYHHLRPEILKFTLHDHVRREAAQALRLGRLVNGSQ 190  
DB 132 WMEERYRHIPGHEVVRTGPCSWTFYHHLRPEILKFTLHDHVRREAAQALRLGRLVNGSQ 191  
QY 181 PSTFVGHVHVRGGDYVHVMPPVWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240  
DB 192 PSTFVGHVHVRGGDYVHVMPPVWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 251  
QY 241 NINASRGDGVVFAAGNIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLAGGDTIYLANYYL 300  
DB 252 NIDTSRSDVVFAAGNIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLAGGDTIYLANYYL 311  
QY 301 PDSFFLKVPKPEAAFLPEWVGIPADLSPLIK 331  
DB 312 PDSFFLKVPKPEAAFLPEWVGIPADLSPLIK 342

Search completed: May 27, 2003, 15:10:45  
Job time : 39.17 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 10:20:58 ; Search time 2880.28 seconds  
(without alignments)  
10791.256 Million cell updates/sec

Title: US-10-040-863-9

Perfect score: 1068

Sequence: 1 ctccagcaggaatagtaa.....gaagagccttcgatgggaa 1068

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vt.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vt.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rnd.\*

36: em\_htg\_man.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1068	100.0	1068	10	AF042743 Rattus no
2	1068	100.0	1149	10	AF264005 Rattus no
3	1054.4	98.7	1847	10	AF131236 Rattus no
4	1052.8	98.6	2984	10	AB006138 Rattus no
5	1040.2	97.4	150840	2	AC121211 Rattus no
6	857.2	80.3	6762	10	AF214656 Mus muscu
7	857.2	80.3	206146	2	AC073774 Mus muscu
8	791.2	74.1	1044	10	AF064792 Mus muscu
9	784.6	71.6	963	10	AB039114 Mus muscu
10	784.6	71.6	963	10	AB039115 Mus muscu
11	784.6	71.6	963	10	AB039116 Mus muscu
12	784.6	71.6	963	10	AB039117 Mus muscu
13	784.6	71.6	963	10	AB039118 Mus muscu
14	784.6	71.6	963	10	AB039119 Mus muscu
15	784.6	71.6	963	10	AB039120 Mus muscu
16	784.6	71.6	963	10	AB039121 Mus muscu
17	784.6	71.6	963	10	AB039122 Mus muscu
18	784.6	71.6	963	10	AB039123 Mus spici
19	669.6	62.7	1555	10	AF131239 Rattus no
20	659.6	61.8	1107	10	AB039213 Mus spici
21	658	61.6	1107	10	AB039209 Mus muscu
22	654.8	61.3	1107	10	AB039212 Mus muscu
23	654.8	61.3	1107	10	AF113532 Mus muscu
24	654.8	61.3	1589	10	MMSEC1 Y09852 M.musculus
25	653.2	61.2	1107	10	AB039206 Mus muscu
26	653.2	61.2	1107	10	AB039207 Mus muscu
27	653.2	61.2	1107	10	AB039208 Mus muscu
28	653.2	61.2	1107	10	AB039211 Mus muscu
29	653.2	61.2	4705	10	AF214657 Mus muscu
30	651.6	61.0	1107	10	AB039210 Mus muscu
31	650	60.9	1107	10	AB039210 Mus muscu
32	650	60.9	1107	10	AB039204 Mus muscu
33	625.6	58.6	5357	10	AB039205 Mus muscu
34	614.8	57.6	2315	4	AF136895 Mus muscu
35	614.8	57.6	2922	4	SSU70881 AF136895 Sus scrofa
36	612.8	57.4	1043	6	AR212339 Sequence
37	612.8	57.4	1043	6	AR212339 Sequence
38	612.8	57.4	1044	4	AF027304 Sus scrofa
39	612.2	57.3	1317	4	OCA12PT X91269 O.cuniculus
40	611.2	57.2	1043	6	AX029101 Sequence
41	594.4	55.7	1062	9	AB015635 Gorilla g
42	592.8	55.5	1062	9	AF080606 Gorilla g
43	592	55.4	1738	4	BF0742GEN X99620 Bos taurus
44	591.2	55.4	1062	9	D87934 Chlorocebus
45	591	55.3	119631	9	AC008888 Homo sapi

#### ALIGNMENTS

RESULT 1  
AF042743 AF042743 1068 bp mRNA linear ROD 01-MAY-2001  
LOCUS Rattus norvegicus alpha 1,2 fucosyltransferase mRNA, partial cds.  
DEFINITION AF042743  
ACCESSION AF042743  
VERSION AF042743.1 GI:2801824  
KEYWORDS Rattus norvegicus.  
SOURCE Rattus norvegicus.  
ORGANISM Rattus norvegicus.  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 1068)  
AUTHORS Sherwood,A.L. and Holmes,E.H.  
TITLE Cloning and expression of the catalytic domain from rat hepatoma



```
/db_xref="GI:8099677"
/translation="MASAQVPSFPLAHLIEFVFTSTIIHLQQRIVKLOPLSEKELP
MTQSSNGTESPEMRDSEOHNGELRCGMFTINSIGRLGNOMGEVATLFAALRNKR
LAFIPASHMNALEPIFRISLPLVHSDTAKKIPQNYHLNDWEERYRHIPGFVPTG
YFCSTFWHLHLEPEILKETLHDHVEDRAQAFRLGLRVNGSOPSTFVGVHVRGDIYV
VPMNVKGVADRGYLEKALDMFRARYSSPVFTVSNMWCNREINASRGDVFAGN
GIEGSPAKDFALLTQCNTIMTIGTFGIWAAYLAGSDTIYLANITLPSDFPLKVPKE
AAFLPEWVGIPADLSPLKALPAPCRSHFHLKAKGVTCYVAGRAF"
BASE COUNT 247 a 341 c 311 g 250 t
ORIGIN
Query Match 100.0%; Score 1068; DB 10; Length 1149;
Best Local Similarity 100.0%; Pred. No. 2.9e-270;
Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTCACGACGATAGTAGCTCAACCCCTGTCTAGAGAGGAANTACCGATGACGACT 60
DB 82 CTCACGACGATAGTAGCTCAACCCCTGTCTAGAGAGGAANTACCGATGACGACT 141
QY 61 CAATGTCTCGGGAACACAGAAAGCCAGAGATGCGACGGGACAGCGAGCATGGG 120
DB 142 CAATGTCTCGGGAACACAGAAAGCCAGAGATGCGACGGGACAGCGAGCATGGG 201
QY 121 AATGGAGACTCGGGGCATGTTCAGGATCAATTCATTGCGCGGCTGGGAACAGATG 180
DB 202 AATGGAGACTCGGGGCATGTTCAGGATCAATTCATTGCGCGGCTGGGAACAGATG 261
QY 181 GCGAATACGCCACACTCTTTCACCTGGCCAGGATGAACGGACGGCTTCGCTCATCCCC 240
DB 262 GCGAATACGCCACACTCTTTCACCTGGCCAGGATGAACGGACGGCTTCGCTCATCCCC 321
QY 241 GCATCCATCCAAACGCTCTAGCGCCCACTTCAGGATCAGCTCCCGGTGTACACAGC 300
DB 322 GCATCCATCCAAACGCTCTAGCGCCCACTTCAGGATCAGCTCCCGGTGTACACAGC 381
QY 301 GACAGGCCAAAGATGCCATGGCAGATTTACCATCTCAAGACTGTGATGAGGAGCGT 360
DB 382 GACAGGCCAAAGATGCCATGGCAGATTTACCATCTCAAGACTGTGATGAGGAGCGT 441
QY 361 TACCGCCACATTCGGGACACTTTGTGCGCTTACAGGGATACCGCTGCCTGGACCTTC 420
DB 442 TACCGCCACATTCGGGACACTTTGTGCGCTTACAGGGATACCGCTGCCTGGACCTTC 501
QY 421 TACACACCTCGGCCACAGATCTCTGAAGAGTTCACCTGCATGACACAGCTGGGGAG 480
DB 502 TACACACCTCGGCCACAGATCTCTGAAGAGTTCACCTGCATGACACAGCTGGGGAG 561
QY 481 GAGGCCAGGCTTCCTCGGTGCTCTGGGTGAATGGAGCAGCCAGCTACTTTTGTG 540
DB 562 GAGGCCAGGCTTCCTCGGTGCTCTGGGTGAATGGAGCAGCCAGCTACTTTTGTG 621
QY 541 GGTGTCCATGTGCGCGGAGGACTATGTGATGTTCATGCTTAATGTGTGAAGGCGGTG 600
DB 622 GGTGTCCATGTGCGCGGAGGACTATGTGATGTTCATGCTTAATGTGTGAAGGCGGTG 681
QY 601 GTGGCTGACCGGGTTACTTGAAAGGCCCTGGATATTCGGGACGCTATTATCT 660
DB 682 GTGGCTGACCGGGTTACTTGAAAGGCCCTGGATATTCGGGACGCTATTATCT 741
QY 661 CCAGTCTTGTGTTACAGCAACGGTATGCGCTGTCGGGGAGCAACATTAATGCTTC 720
DB 742 CCAGTCTTGTGTTACAGCAACGGTATGCGCTGTCGGGGAGCAACATTAATGCTTC 801
QY 721 CGAGGAGACGTGGTGTTCGCGGGCAATGTATGTAGGGGTGCGCCAGCCAAAGACTTCGG 780
DB 802 CGAGGAGACGTGGTGTTCGCGGGCAATGTATGTAGGGGTGCGCCAGCCAAAGACTTCGG 861
QY 781 CTGCTCAOCCAGTGAACACACCATCATCTACTATTTGGACCTTTGGGATTTGGGCTGCC 840
DB 862 CTGCTCAOCCAGTGAACACACCATCATCTACTATTTGGACCTTTGGGATTTGGGCTGCC 921
QY 841 TACCTGGAGGTGGTATACCATCTACTATTTAGCCACTACACCTTCGCGATTTCCGCTC 900
DB 922 TACCTGGAGGTGGTATACCATCTACTATTTAGCCACTACACCTTCGCGATTTCCGCTC 981
QY 901 CTCAAAGTCTTTAAGCAGAGGAGCGCTTCCTACCGGAATGGTGGGCATCCCTGCCGAT 960
DB 982 CTCAAAGTCTTTAAGCAGAGGAGCGCTTCCTACCGGAATGGTGGGCATCCCTGCCGAT 1041
QY 961 CTGTCCCACTCTCTAAGGATTAAACACAGCGCTGTCTCGTGGTCCCACTTCCACCTCAAG 1020
DB 1042 CTGTCCCACTCTCTAAGGATTAAACACAGCGCTGTCTCGTGGTCCCACTTCCACCTCAAG 1101
QY 1021 GCAAAAGGAGTCACTTGTACGTGCGAGGAGAGCGCTTCGTGATGGAA 1068
DB 1102 GCAAAAGGAGTCACTTGTACGTGCGAGGAGAGCGCTTCGTGATGGAA 1149
```

```
RESULT 3
AF131238 1847 bp DNA linear ROD 08-MAR-2001
LOCUS Rattus norvegicus alpha 1,2-fucosyltransferase B (FTB) gene,
DEFINITION complete cds.
ACCESSION AF131238
VERSION AF131238.1 GI:4580672
KEYWORDS Rattus norvegicus.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1847)
AUTHORS Bureau,V., Marionneau,S., Cailletau-Thomas,A., Le Mouillac-Vaidye,B.,
Lieber,T. and Le Pendu,J.
TITLE Comparison of the three rat GDP-L-fucose:beta-D-galactoside
2-alpha-L-fucosyltransferases, FTa, FTB and FTc
JOURNAL Eur J. Biochem. 268 (4), 1006-1019 (2001)
MEDLINE 21099374
PUBMED 11179967
REFERENCE 2 (bases 1 to 1847)
AUTHORS Bureau,V., Le Mouillac-Vaidye,B., Liehr,T., Denis,M. and Le Pendu,J.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-1999) INSERM U419, Institut de Biologie, 9 Quai
Moncoussu, Nantes 44035, France
FEATURES
Source
1..1847
/organism="Rattus norvegicus"
/strain="BDIX"
/db_xref="taxon:10116"
/chromosome="1"
/map="1q22-q31"
<211..>1275
/gene="FTB"
<211..>1275
/gene="FTB"
/product="alpha 1,2-fucosyltransferase B"
211..1275
/gene="FTB"
/notes="glycosyltransferase; alpha 2-fucosyltransferase B"
/codon_start=1
/product="alpha 1,2-fucosyltransferase B"
/protein_id="AAD24469.1"
/db_xref="GI:4580673"
/translation="MASAQVPSFPLAHLIEFVFTSTIIHLQQRIVKLOPLSEKELP
MTQSSNGTESPEMRDSEOHNGELRCGMFTINSIGRLGNOMGEVATLFAALRNKR
LAFIPASHMNALEPIFRISLPLVHSDTAKKIPQNYHLNDWEERYRHIPGFVPTG
YFCSTFWHLHLEPEILKETLHDHVEDRAQAFRLGLRVNGSOPSTFVGVHVRGDIYV
VPMNVKGVADRGYLEKALDMFRARYSSPVFTVSNMWCNREINASRGDVFAGN
GIEGSPAKDFALLTQCNTIMTIGTFGIWAAYLAGSDTIYLANITLPSDFPLKVPKE
AAFLPEWVGIPADLSPLKALPAPCRSHFHLKAKGVTCYVAGRAF"
BASE COUNT 452 a 503 c 457 g 435 t
ORIGIN
Query Match 98.7%; Score 1054.4; DB 10; Length 1847;
Best Local Similarity 99.8%; Pred. No. 1.1e-266;
Matches 1066; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
```

QY	1	CTCCAGCAGGANTAGTGAAGCTCCACCCCTGTACAGAGAGGAATTACCGATGACGACT	60
Db	292	CTCCAGCAGGGAATAGTGAAGCTCCAAACCCCTGTACAGAGAAGGAATTACCGATGACGACT	351
QY	61	CAAAATGCTTCGGGAAACACAGAAAGCCCAAGAGATGCGACGGACACGCGAGCAGCATGGG	120
Db	352	CAAAATGCTCTCGGAAACACAGAAACCCCAAGAGATGCGACGGACACGCGAGCAGCATGGG	411
QY	121	ANTGGAGAGCTGGGGGCATGTTACAGATCAAATCCATTTGCCCGGCTGGGGAACCCAGATG	180
Db	412	AATGGAGAGCTGGGGGCATGTTACAGATCAAATCCATTTGCCCGGCTGGGGAACCCAGATG	471
QY	181	GGCGAATACCGCACACATCTTTGCACATGGCCAGATGAACGAGCGGTTGGTTTCATCCCC	240
Db	472	GGCGAATACCGCACACATCTTTGCACATGGCCAGATGAACGAGCGGTTGGTTTCATCCCC	531
QY	241	GCATCCATGCACAAACGCTCTTAGCGCCCATCTTCAGAGTACAGCTCCCGGTGTTACACAGC	300
Db	532	GCATCCATGCACAAACGCTCTTAGCGCCCATCTTCAGAGTACAGCTCCCGGTGTTACACAGC	591
QY	301	GACACGGCCAAAAGATCCCATGGGAGAAATTACCATCTCAACGACTGGATGGAGGAGCGT	360
Db	592	GACACGGCCAAAAGATCCCATGGGAGAAATTACCATCTCAACGACTGGATGGAGGAGCGT	651
QY	361	TACCGCCACATTCGGGACACTTTGTGCGGTTTCACGGGATACCCGCTCTCCTCGACCTTC	420
Db	652	TACCGCCACATTCGGGACACTTTGTGCGGTTTCACGGGATACCCGCTCTCCTCGACCTTC	711
QY	421	TACACACCTGGCGCCAGAGATCTCGAAGAGATTCAACCTGATCAACACGCTGCGGGAG	480
Db	712	TACACACCTGGCGCCAGAGATCTCGAAGAGATTCAACCTGATCAACACGCTGCGGGAG	771
QY	481	GAGGCCACAGCCCTCTGGGTGCTGCGGGTGAATGGGAGCCAGCGAGTACTTTGTG	540
Db	772	GAGGCCACAGCCCTCTGGGTGCTGCGGGTGAATGGGAGCCAGCGAGTACTTTGTG	831
QY	541	GGTGTCCATGTGGCCGAGGGACTATGTGCATGTTCATGCTTAATGTGGAAGGCGGTG	600
Db	832	GGTGTCCATGTGGCCGAGGGACTATGTGCATGTTCATGCTTAATGTGGAAGGCGGTG	891
QY	601	GTGGCTACCGGGTTACCTGGAAGAGCCCTCGATCTTCGGGCACGCTATTTCATCT	660
Db	892	GTGGCTACCGGGTTACCTGGAAGAGCCCTCGATCTTCGGGCACGCTATTTCATCT	951
QY	661	CCAGTCTTCGTGTTACAAGCAACGGTATGGCCCTGGTGGCGGAGAACATTAATGCTTCC	720
Db	952	CCAGTCTTCGTGTTACAAGCAACGGTATGGCCCTGGTGGCGGAGAACATTAATGCTTCC	1011
QY	721	CGAGGAGAGCTGGTGTCTCGGGCAATGTTATGAGGGGTGCGCCAGCAGGACTTCGGG	780
Db	1012	CGAGGAGAGCTGGTGTCTCGGGCAATGTTATGAGGGGTGCGCCAGCAGGACTTCGGG	1071
QY	781	CTGCTCACCCAGTGCACACACCATCATGACATATTGGGACCTTTGGGATTTGGGCTGCC	840
Db	1072	CTGCTCACCCAGTGCACACACCATCATGACATATTGGGACCTTTGGGATTTGGGCTGCC	1131
QY	841	TACCTGGCAGGTGGTATACCATCTACTTAGCCAACTACACCTTCGGGATTTCCGGTTC	900
Db	1132	TACCTGGCAGGTGGTATACCATCTACTTAGCCAACTACACCTTCGGGATTTCCGGTTC	1191
QY	901	CTCAAAAGTCTTTAAGCCAGAGCAGCGCTTCCTACCCGAATGGTGGGCACTCCCTGCCGAT	960
Db	1192	CTCAAAAGTCTTTAAGCCAGAGCAGCGCTTCCTACCCGAATGGTGGGCACTCCCTGCCGAT	1251
QY	961	CTGTCCCCACTGCTTAAGGCATTAACACCAAGCCTGTCTCGTGGTCCCACTTCCACCTCAAG	1020
Db	1252	CTGTCCCCACTGCTTAAGGCATTAACACCAAGCCTGTCTCGTGGTCCCACTTCCACCTCAAG	1310
QY	1021	GCAAAAGGAGTCACTTGTACGTTCGGCAGNAGAGCCTTCTGTATGGGAA	1068
Db	1311	GCAAAAGGAGTCACTTGTACGTTCGGCAGNAGAGCCTTCTGTATGGGAA	1358

```

RESULT 4
AB006138          2984 bp      mRNA       linear   ROD 05-FEB-1999
LOCUS             Rattus norvegicus FTB mRNA for alpha 1,2-fucosyltransferase,
DEFINITION        complete cds.
ACCESSION         AB006138
VERSION           AB006138.1 GI:2317265
KEYWORDS          FTB; alpha 1,2-fucosyltransferase.
SOURCE            Rattus norvegicus colon cancer cell_line:RCN-9 cDNA to mRNA.
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                  Rattus.
REFERENCE         1 (sites)
AUTHORS           Soejima,M., Wang,B., Koda,Y. and Kimura,H.
TITLE             Two distinct rat Gbp-L-fucose-B-D-galactoside
                  2-a-L-fucosyltransferase genes
JOURNAL           Unpublished
REFERENCE         2 (bases 1 to 2984)
AUTHORS           Koda,Y.
TITLE             Direct Submission
JOURNAL           Submitted (04-AUG-1997) Yoshiro Koda, Kurume University, School of
                  Medicine, Department of Forensic Medicine; Asahimachi 67, Kurume,
                  Fukuoka 830, Japan (E-mail:ykoda@med.kurume-u.ac.jp,
                  Tel:0942-31-7554, Fax:0942-31-7700)
FEATURES          Location/Qualifiers
                   1..2384
                       /organism="Rattus norvegicus"
                       /db_xref="taxon:10116"
                       /chromosome="1"
                       /map="lq22-q31"
                       /cell_line="RCN-9"
                       /tissue_type="colon cancer"
                   1..2384
                       /genes="FTB"
                       /gene="FTB"
                       /codon_start=1
                       /product="alpha 1,2-fucosyltransferase"
                       /protein_id="BAA21742.1"
                       /db_xref="GI:2317266"
                       /translation="MASAQPFPSFLAHFLIFVFTVTTHLQQRIVKLQPLSERKELP
MTQMSSGATETSEMERDSEOHNGELRGMTINSIGRLGNOMGEVATIFALRWKR
LARLPASMINALAPIRESILPLVHSDTKKIPWNYHLNDWEERYRHITGFHVRTG
IPCSYTFHHLEPEIKLETLDHVEEAQAFLURLKVNGSQSPFTFVGVVHVRGDIYV
VMPNVKWGVADREYILEKALDMFRARYSSPVFYVTSNGMAWCRENINSGRDVVFAGN
GIEGSPAKFALLTCQNHITMITGTGTGAAYLAGGDITLYLANITLPDSPFLKVKEPKE
AAPLPWVGVIPADLSPLLKH"
BASE COUNT       729 a    808 c    761 g    686 t
ORIGIN
Query Match.          98.5%; Score 1052.8; DB 10; Length 2984;
Best local similarity 99.7%; Pred.No.2.9e-266;
Matches 1065; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 CTCACAGCGGTAAGTCTGAAGCTCCAAACCCTCCTCAGAGAAGGAATTACCGATGACGACT 60
|||||
DB 295 CTCACAGCGGAATAGTAGCTGCCAACCCCTCAGAGAAGGAATTACCGATGACGACT 354
|||||

QY 61 CAATAATGCCTCGGGAAACACAGAAAACCCAGAGATCGCACGGGACACGACGACGATGGG 120
|||||
DB 355 CAAATGTCCTCGGGAAACACAGAAAGCCAGAGATCGCACGGGACACGACGACGATGGG 414
|||||

QY 121 AATGCGAGAGTCGGGGCATGTTACAGATCAATTCATTGGCGGGCTGGGGACACAGATG 180
|||||
DB 415 AATGCGAGAGTCGGGGCATGTTACAGATCAATTCATTGGCGGGCTGGGGACACAGATG 474
|||||

QY 181 GGCGAATACGCCACACTCTTGACACTGGCCAGGATGAACGAGCGGCTCGCGTTCATCCCC 240
|||||
DB 475 GCGGAATACGCCACACTCTTGACACTGGCCAGGATGAACGAGCGGCTCGCGTTCATCCCC 534
|||||

QY 241 GCATCATGACACAAC3CTCTAGCGCCCATCTTCAGGATCAGCCTCCCGGTGTTACACAGC 300

```

```
Db      535 GCATCCATGCACACGCTCTAGCCCATCTTCAGGATCAGCCTCCCGGTATACACAGC 594
Qy      301 GACACGCCAAAAGATCCCATCGCAGATACATCTCAACGACTGGATGGAGGACGCT 360
Db      595 GACACGCCAAAAGATCCCATCGCAGATACATCTCAACGACTGGATGGAGGACGCT 654
Qy      361 TACCGCCACATTCGCGGACACTTTGTCGGTTCACGGATACCGTGCTCTCGTGAACCTTC 420
Db      655 TACGCCACATTCGCGGACACTTTGTCGGTTCACGGATACCGTGCTCTCGTGAACCTTC 714
Qy      421 TACACACACCTGCGCCGACAGATCTCTGAAGAGTTCACCTGCAATGACACCGTCCGGAG 480
Db      715 TACACACACCTGCGCCGACAGATCTCTGAAGAGTTCACCTGCAATGACACCGTCCGGAG 774
Qy      481 GAGGCCACAGCCTTCTCGGTCGGTCTCGGTCGAGTGAATGGGAGCCAGCAGTACTTTTGTG 540
Db      775 GAGGCCACAGCCTTCTCGGTCGGTCTCGGTCGAGTGAATGGGAGCCAGCAGTACTTTTGTG 834
Qy      541 GGTGTCCATGTGCGCCGAGGGGACTATGTGCATGTCAATGCTATGTGTGGAAGGGCGTG 600
Db      835 GGTGTCCATGTGCGCCGAGGGGACTATGTGCATGTCAATGCTATGTGTGGAAGGGCGTG 894
Qy      601 GTGCTCATCCGCGGTTACCTTGGAAAGCCCTGGATATGTTCCGGGACGCTATTCATCT 660
Db      895 GTGCTCATCCGCGGTTACCTTGGAAAGCCCTGGATATGTTCCGGGACGCTATTCATCT 954
Qy      561 CCATGCTCTGTGTTACAGCAACGGTATGCTCTGTCGCGGGAGACATTAATGCTTCC 720
Db      955 CCATGCTCTGTGTTACAGCAACGGTATGCTCTGTCGCGGGAGACATTAATGCTTCC 1014
Qy      721 CGAGGACGCTGTGTTGCGGGGCAATGGTATTCAGGGGTGCCAGCAAGGACTTCCGC 780
Db      1015 CGAGGACGCTGTGTTGCGGGGCAATGGTATTCAGGGGTGCCAGCAAGGACTTCCGC 1074
Qy      781 CTGCTCAACCCAGTGCACACACCATCATGACTATTTGGGACCTTTGGGATTTGGGCTGCC 840
Db      1075 CTGCTCAACCCAGTGCACACACCATCATGACTATTTGGGACCTTTGGGATTTGGGCTGCC 1134
Qy      841 TACCTGGCAGGTGGTATACCATCTACTTACCACTACACCTTCCGGATTCCTCGTTC 900
Db      1135 TACCTGGCAGGTGGTATACCATCTACTTACCACTACACCTTCCGGATTCCTCGTTC 1194
Qy      901 CTCAAAGTCCTTAAAGCCAGAGCGACCTTCCTACCGAATGGTGGGATCCCTCCGAT 960
Db      1195 CTCAAAGTCCTTAAAGCCAGAGCGACCTTCCTACCGAATGGTGGGATCCCTCCGAT 1254
Qy      961 CTGTCGCCACCTTCCTTAAGGCAATTAACACAGCCTGCTCGGTGCCATTCACCTCAAG 1020
Db      1255 CTGTCGCCACCTTCCTTAAGGCAATTAACACAGCCTGCTCGGTGCCATTCACCTCAAG 1313
Qy      1021 GCAAAAGGAGTCNCTTGTAGTTCGACAGAGGCTTCTGATGGAA 1068
Db      1314 GCAAAAGGAGTCNCTTGTAGTTCGACAGAGGCTTCTGATGGAA 1361

RESULT 5
AC121211
LOCUS      150840 bp      DNA      linear      HTG 24-JUL-2002
DEFINITION Rattus norvegicus clone CH230-309A14, *** SEQUENCING IN PROGRESS
            *** 63 unordered pieces.
ACCESSION  AC121211
VERSION     AC121211.2  GI:21909463
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE     Rattus norvegicus.
            Rattus norvegicus.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 150840)
            Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-olsman, F.R., Allen, C.,
            Alsbrooks, S.L., Amaral, H.C., Are, J.R., Ayele, M., Banks, T.,
            Barbara, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D.,
```

```
Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hoques, M., Holloway, C., Hollins, B.,
Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlsone, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louissegh, H.,
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okunolu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,
Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H., Tang, H.,
Tanton, J., Tatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tatney, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Unpublished
2 (bases 1 to 150840)
Worley, K.C.
Direct Submission
Submitted (16-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 150840)
Worley, K.C.
Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 19, 2002 this sequence version replaced gi:20806241.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GYKK
Center clone name: CH230-309A14
----- Summary Statistics
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 106837 bases at least Q40
Consensus quality: 113151 bases at least Q30
Consensus quality: 116417 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 63 contigs. The true order of the pieces
* is not known and their order in this sequence record is
```





Db	143184	CTACACACCTCGGCCAGAGATCTGAAGAGTTCACCTCATGACCACCTGCGGA	143243	
QY	480	GGAGGCCAGGCTTCTCGTGGTCTGGGGTGAATGGAGCCAGCCGAGTACTTTGT	539	
Db	143244	GGAGGCCAGGCTTCTCGTGGTCTGGGGTGAATGGAGCCAGCCGAGTACTTTGT	143303	
QY	540	GGGTGTCCATGTGCGCGAGGGGACTATGTGCATGTGATGCCCTTAATGTGTGAAGGCGT	599	
Db	143304	GGGTGTCCATGTGCGCGAGGGGACTATGTGCATGTGATGCCCTTAATGTGTGAAGGCGT	143363	
QY	600	GTGGGTGACCGGGGTACTGGAAGGCCCTGGATGTTCCGGCAGCTATTCAATC	659	
Db	143364	GTGGGTGACCGGGGTACTGGAAGGCCCTGGATGTTCCGGCAGCTATTCAATC	143423	
QY	660	TCCAGTCTTGGTGTACAGCAACGGTATGSCCTGGTCCGGGAGAACATTAATGCTTC	719	
Db	143424	TCCAGTCTTGGTGTACAGCAACGGTATGSCCTGGTCCGGGAGAACATTAATGCTTC	143483	
QY	720	CCGAGGAGCTGGTCTTCGGGGCAATGGTATGAGGGGTGCGCCAGCAAGACTTCGC	779	
Db	143484	CCGAGGAGCTGGTCTTCGGGGCAATGGTATGAGGGGTGCGCCAGCAAGACTTCGC	143543	
QY	780	GCTGTTCACCCAGTGAACACACCATCATGACTATTGGGACCTTTGGATTGGCTGC	839	
Db	143544	GCTGTTCACCCAGTGAACACACCATCATGACTATTGGGACCTTTGGATTGGCTGC	143603	
QY	840	CTACCTGGCAGTGTGATPACCATCTACTATGCCAATACACCTTCCGGATTCCTCGTT	899	
Db	143604	CTACCTGGCAGTGTGATPACCATCTACTATGCCAATACACCTTCCGGATTCCTCGTT	143663	
QY	900	CCTCAAGTCTTTAAGCCAGAGGAGCTTCTACCCGAATGGTGGGATCCCTGCCGA	959	
Db	143664	CCTCAAGTCTTTAAGCCAGAGGAGCTTCTACCCGAATGGTGGGATCCCTGCCGA	143723	
QY	960	TCGTCCCACTCCCTTAAGGCATTAAACACGCTTCTCGTCCCTCCACTCCACTCAA	1019	
Db	143724	TCGTCCCACTCCCTTAAGGCATTAAACACGCTTCTCGTCCCTCCACTCCACTCAA	143782	
QY	1020	GCRAAAGAGTCACTTGTAGTGCAGGAAGCCCTTCTGATGGAA	1068	
Db	143783	GCRAAAGAGTCACTTGTAGTGCAGGAAGCCCTTCTGATGGAA	143831	
RESULT 6				
LOCUS	AF214656			
DEFINITION	Mus musculus alpha(1,2)fucosyltransferase FUT2 (Fut2) gene, complete cds.	6762 bp	DNA	linear
ACCESSION	AF214656			
VERSION	AF214656.1	GI:7288504		
KEYWORDS				
SOURCE	Mus musculus.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Domino, S.B., Zhang, L. and Lowe, J.B.			
TITLE	1 (bases 1 to 6762) Molecular cloning, genomic mapping, and expression of two secretor blood group alpha (1,2)fucosyltransferase genes differentially regulated in mouse uterine epithelium and gastrointestinal tract			
JOURNAL	J. Biol. Chem. 276 (26), 23748-23756 (2001)			
MEDLINE	21316345			
PubMed	11323419			
REFERENCE	2 (bases 1 to 6762) Domino, S.B. and Lowe, J.B.			
AUTHORS	Mus musculus alpha(1,2)fucosyltransferase FUT2 (Sec2) genomic sequence			
TITLE	Unpublished			
JOURNAL	3 (bases 1 to 6762) Domino, S.B. and Lowe, J.B.			
REFERENCE	Direct Submission			
AUTHORS	Submitted (13-DEC-1999) HHMI, University of Michigan, 1150 W. Med.			
TITLE				
JOURNAL				

FEATURES	Cur. Dr., Ann Arbor, MI 48109-0650, USA
SOURCE	1. 6762
	/organism="Mus musculus"
	/strain="129/Ola"
	/db_xref="taxon:10090"
	/chromosome="7"
	/map="23.2; between K1k1 and Ldh3"
	/cell_line="E14"
	/cell_type="ES"
gene	<2210..>3253
	/genes="Fut2"
	/note="Sec2"
mRNA	<2210..>3253
	/genes="Fut2"
	/product="alpha(1,2)fucosyltransferase FUT2"
CDS	2210..3253
	/gene="Fut2"
	/EC_number="2.4.1.69"
	/note="fucosyltransferase; similar to the Mus musculus strain ICR Sec2 sequence deposited at GenBank Accession Number AF064792; similar to the Homo sapiens secretor blood group alpha(1,2)fucosyltransferase FUT2 secretor deposited at GenBank Accession Numbers D87942 and U17894"
	/codon_start=1
	/product="alpha(1,2)fucosyltransferase FUT2"
	/protein_id="AAF45146.1"
	/db_xref="GI:7288505"
	/translation="MASAQPFSPFLAHLFVFTSTIHLQORIVKLTISEKLQAVOMSPNARDTMOQSAKIGFTINSIGRLGNOMGEYAILFALAMNGRLAFIPESMENLAPTRISLPHSDTARRIPWYHLNDWMEERYRHIPGOVYRTGYPCSKTFYHLRPEILKEFTLHDHVRERACATLRLGVNSOPSTFVGVHVRGDIVVHVPKWKGVVADRGYLEKALDRERARYSPFVVTNSGMAWRENINTSLGDDVVFAGNIGESPAKIPADLSTPLLRH"
	polyA_signal
	4943..4948
	polyA_signal
	6333..6338
BASE COUNT	1738 a 1682 c 1647 g 1691 t 4 others
ORIGIN	
Query Match	80.3%; Score 857.2; DB 10; Length 6762;
Best Local Similarity	89.3%; Pred. No. 9.2e-215;
Matches	954; Conservative 0; Mismatches 93; Indels 21; Gaps 2;
QY	1 CTCCAGGCGGGAATAGTGAAGCTCCAAACCCCTGTCCAGAGAAGAAATACCGATGAGGACT 60
Db	2291 CTCCAGGCGGGAATAGTGAAGCTCCAAACCCCTGTCCAGAGAAGAAATAC---AGGCGGT 2347
QY	61 CAAATGTCTCTCGGGAACACAGAAAGCCCGGACGAGATGCGGACGAGCAGCAGCATGG 120
Db	2348 CAAATGTCTCTCACCAACGCGCGCAGACACATCCAGCA-----G 2389
QY	121 ATGAGAGCTCGGGGATGTTACGATCAATTCATTCGCGCGCTGGGGAACAGATG 180
Db	2390 ATGCGCAAGCTCGAGGCAATTCAGATCAATTCATTCGCGCGCTGGGGAACAGATG 2449
QY	181 GCGAATAGCCACACTCTTTCACATGCGCCAGATGACGAGCGCTTCGCTTCATCC 240
Db	2450 GCGAATAGCTACATGTTTTCACATGCGCCAGATGACGAGCGCTTCGCTTCATCC 2509
QY	241 GCATCCATGACACGCTCTAGCGCCCATCTTCAGGATCAGCTCCCGGTGTACACAGC 300
Db	2510 GAATCCATGACACGCTCTAGCGCCCATCTTCAGGATCAGCTCCCGGTGTACACAGC 2569
QY	301 GACAGGCCAAAGATCCCATGGCAGATTAACCATCTCAACGATGGATGGAGAGCGT 360
Db	2570 GACAGCCAGAAGATCCCGTGGCAGATTAACCATCTCAACGATGGATGGAGAGCGT 2629
QY	361 TACGCCACATTCGGGACACTTTTGGCTTCAGGATACCGGTGCTCTCTGGACCTTC 420
Db	2630 TACGCCACATTCGGGACACTTTTGGCTTCAGGATACCGGTGCTCTCTGGACCTTC 2689
QY	421 TACCACCACTTCGCCCAAGAGATCCTGAAGGAGTTCACCTGCTGATGACCACTGCGGAG 480



Db	13777	GGGCAATATGCTACATGTTGGTGCAGTGGCCAGGATGAACGCTGGCTTGCCTTCATCCCT	13718
Qy	241	GCATCATGACACACGCTTAGCGCCCATCTCAGGATCAGCCTCCGGTGTACACAGC	300
Db	13717	GAATCCATGACACACGCTTAGCGCCCATCTCAGGATCAGCCTCCGGTGTACACAGC	13658
Qy	301	GACAGCCGAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGCGT	360
Db	13657	GACACAGCAGAAAGATCCCGTGGCAGAAATACCATCTCAACGACTGGATGGAGGCGT	13598
Qy	361	TACGGCCACATTCGGGACACTTGTGCGCTTCAACGGATACCGTGCTCTGGACCTC	420
Db	13597	TACGGCCACATTCGGGACACTTGTGCGCTTCAACGGATACCGTGCTCTGGACCTC	13538
Qy	421	TACACACCTGGCCGACAGATCTCAAGGATTCACCTGTCATGACACGCTGCGGGAG	480
Db	13537	TACACACCTGGCCGACAGATCTCAAGGATTCACCTGTCATGACACGCTGCGGGAG	13478
Qy	481	GAGGCCAGCGCTTCCTGCTGCTGCTGCGGTGAATGGGAGCGAGTACCTTTTGTG	540
Db	13477	GAGGCCAGCGCTTCCTGCTGCTGCTGCGGTGAATGGGAGCGAGTACCTTTTGTG	13418
Qy	541	GGTGTCCATGTGGCGGAGGACTATGTCATGTCATGCTCATGTCGAGGGCGTG	600
Db	13417	GGTGTCCATGTGGCGGAGGACTATGTCATGTCATGCTCATGTCGAGGGCGTG	13358
Qy	601	GTGGCTGACCGGGGTACCTGGAAAAGCGCTGGATATGTCGGGCACGCTATTCATCT	660
Db	13357	GTGGCTGACCGGGGTACCTGGAAAAGCGCTGGATATGTCGGGCACGCTATTCATCT	13298
Qy	661	CGAGTCTTCTGTTACAGCACAGGTATGGCTGPRGCGGGAGAACATTAATGCTTC	720
Db	13297	CGAGTCTTCTGTTACAGCACAGGTATGGCTGPRGCGGGAGAACATTAATGCTTC	13238
Qy	721	CGAGGAGACGTGTGTGTCGGGCAATGATGATGAGGGTCGCCAGCAGGACTTCGG	780
Db	13237	CTAGGAGACGTGTGTGTCGGGCAATGATGATGAGGGTCGCCAGCAGGACTTCGG	13178
Qy	781	CTGCTCACCAGTGCACACACATCATGACTATTGGGACCTTTGGGATTTGGCTGCC	840
Db	13177	CTGCTCACCAGTGCACACACATCATGACTATTGGGACCTTTGGGATTTGGCTGCC	13118
Qy	841	TACCTGGCAGTGGTGATACCATCTACTTAGCAACTACACCTTCGGGATTCGCCCTC	900
Db	13117	TACCTGGCAGTGGTGATACCATCTACTTAGCAACTACACCTTCGGGATTCGCCCTC	13058
Qy	901	CTCAAGTCTTTAAGCAGAGGAGCTTCCTACCGAATGGTGGCATCCCTGCCGAT	960
Db	13057	CTCAAGTCTTTAAGCAGAGGAGCTTCCTACCGAATGGTGGCATCCCTGCCGAGC	12998
Qy	961	CTGTCCCATCTTAAAGSCATTAAACACGACCTGTCTCGGCCCATCTCCACCTCAAG	1020
Db	12997	CTGTCCCATCTTAAAGSCATTAAACACGACCTGTCTCGGCCCATCTCCACCTCAAG	12938
Qy	1021	GCAAGGAGTCACTGTTACGTCGAGGAGAGCCCTTCGTGATGGGAA	1068
Db	12937	GCAGAGGAGTCACTGTTACGTCGAGGAGAGCCCTTCGTGATGGGAA	12890
RESULT	8		
LOCUS	AF064792		
DEFINITION	Mus musculus GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase mRNA, complete cds.		
ACCESSION	AF064792		
VERSION	AF064792.1	GI:3142705	
KEYWORDS	Mus musculus.		
SOURCE	Mus musculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (bases 1 to 1044)		
REFERENCE	1 (bases 1 to 1044)		
AUTHORS	Lin, B., Hayashi, Y., Saito, M., Sakakibara, Y., Yanagisawa, M. and		

Iwamori, M.		TITLE			
GDP-fucose: beta-galactoside alpha1,2-fucosyltransferase, MFUT-II, and not MFUT-I or -III, is induced in a restricted region of the digestive tract of germ-free mice by host-microbe interactions and cycloheximide		JOURNAL			
Biochim. Biophys. Acta 1487 (2-3), 275-285 (2000)		MEDLINE			
20471982		PUBMED			
11018479		REFERENCE			
2 (bases 1 to 1044)		AUTHORS			
Lin, B., Hayashi, Y., Saito, M., Sakakibara, Y., Yanagisawa, M. and Iwamori, M.		Direct Submission			
Submitted (12-May-1998) Chemistry, Faculty of Science & Technology, Kinki University, Kowakae, Higashi-Osaka 577-8502, Japan		JOURNAL			
Location/Qualifiers		FEATURES			
1..1044		source			
/organism="Mus musculus"		CDS			
/strain="ICR"		1..1044			
/db_xref="taxon:10090"		/EC_number="2.4.1.69"			
/tissue_type="gastrointestinal tract"		/note="sec2; Se"			
1..1044		/codon_start=1			
/product="GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase"		/protein_id="AAC16897.1"			
/db_xref="GI:3142706"		/translation="MASAQVFFSFLAHFLIFVFTSTIIHLOQIVKLTSEKELAVQSSNPAARTMQSAKLGITINSIGRLQNMGEYATLALARMNRLAIPISMHNALPFRISLPVLSHTDARRIPQWYHINDMEERYHRIHQVYRFTGPCSWTFYHLRLPELLEFTLHDHVEEAQFLRLGVNSQSTFVGHVHRGDDYVHVPKYHKGVDRCYLEKALDRPRARYSSPVFVTSNGMWCNENINISLGDVVFAGNGIEGSPAKDPAALTCOHNHITWTGTFGIWAAIYLAGGDTIYLANITLDPSPFLKIFKPAAPFLPEW		BASE COUNT	
229 a		318 c			
271 g		226 t			
ORIGIN		Query Match			
74.1%; Score 791.2; DB 10; Length 1044;		Best Local Similarity			
89.9%; Pred. No. 2.2e-197;		Matches 879; Conservative 0; Mismatches 78; Indels 21; Gaps 2;			
QY	1	CTCCAGCAGCGAATAGTGAAGCTCCAAACCCCTGTCCAGAGGAAGTAATACCGATCACCACCT	60		
DB	82	CTCCAGCAGCAAAATAGTGAAGCTCCAAACCCCTGTCCAGAGGAAGTAATACAGG---CGGTT	138		
QY	61	CAATGTCTCTGGGAAACACAGAAAGCCGACGATGCGACGGGACACGACGACATCGG	120		
DB	139	CAATGTCTCTCACAAACGCGGCAAGAACACATGCAGCA-----G	180		
QY	121	AATGAGAGCTGGGGCATGTTTCACGATCAATTCATTTGCCCGCTGGGGAACCATG	180		
DB	181	AGTCCAGAGCTGACGGGCATTTACGATCAATTCATTCATTCATTCATTCATTCATTC	240		
QY	181	GGGGAATACGCCACATCTTTTGACTGGCCAGGATGAACGAGCGCTTGGCTTCATCCCT	240		
DB	241	GGGGAATATGCTACATTTGTTGACTGGCCAGGATGAACGAGCGCTTGGCTTCATCCCT	300		
QY	241	GCATCCATGCACACGCTCTTAGGCCCATCTTTCAGGATCAGCCCTCCCGTGTACACAGC	300		
DB	301	GAATCCATGCACACGCTCTTAGGCCCATCTTTCAGGATCAGCTCTCCCGTGTACACAGC	360		
QY	301	GACAGGCCCAAAAGATCCCATGCCAGAAATACATCTCAACGACTGGATGGAGGCGT	360		
DB	361	GACAGGCCCAAAAGATCCCGTGGCAGAAATACATCTCAACGACTGGATGGAGGCGT	420		
QY	361	TACCGCCACATTCGGGACACTTTTGTGCGCTTTCACGGATACCGCTCTCTGGACCTTC	420		
DB	421	TACCGCCACATTCGGGACACTTTTGTGCGCTTTCACGGATACCGCTCTCTGGACCTTC	480		
QY	421	TACACACCTGGGCCGACAGATCTCTGAAGAGTTCACCTCTGCATGACACGCTGGGAG	480		
DB	481	TACACACCTGGGCCGACAGATCTCTGAAGAGTTCACCTCTGCATGACACGCTGGGAG	540		



Db 739 CTAGGAGACGTGGTGTTCGGGCAATGGTATTGAGGCTCACCAGCGCAAGGACTTCGCG 798  
QY 781 CTGCTCACCAGTGCACACCAACCATCATGACTATTGGGACTTTGGGATTTGGGCTGCC 840  
Db 799 CTCTCACCAGTGCACACCAACCATCATGACTATTGGGACTTTGGGATTTGGGCTGCC 858  
QY 841 TACCTGGCAGTGGTGATACCATCTACTTACGCAACTACACCTTCCGGATTTCCTCGTTC 900  
Db 859 TACCTGGCAGTGGTGATACCATCTACTTACGCAACTACACCTTCCGGATTTCCTCGTTC 918  
QY 901 CTCGAAGTCTTTAAGCCAGAGGACGCTTCTCCTACCCGAATGGGTG 945  
Db 919 CTCGAAGTCTTTAAGCCAGAGGACGCTTCTCCTCCTCCGAGTGGATG 963

RESULT 10  
AB039115  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
gene  
CDS

AB039115 963 bp DNA linear ROD 02-OCT-2001  
Mus musculus Fut2 gene for GDP-L-fucose:beta-D-galactoside  
2-alpha-L-fucosyltransferase, partial cds, strain:BFM/2Msf.  
AB039115  
AB039115.1 GI:15822975  
FUT2; GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase.  
Mus musculus (sub-species:brevirostris, strain:BFM/2Msf) DNA.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1  
Li,Y., Kitano,T., Koide,T., Shiroishi,T., Moriwaki,K. and  
Saitou,N.  
Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes  
of Five Mus musculus subspecies  
Unpublished  
2 (bases 1 to 963)  
Li,Y. and Saitou,N.  
Direct Submission  
Submitted (29-FEB-2000) Naruya Saitou, National Institute of  
Genetics, Laboratory of Evolutionary Genetics, 1111 Yata, Mishima,  
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,  
URL:http://sayer.lab.nig.ac.jp/, Tel:81-559-81-6790,  
Fax:81-559-81-6789)  
Location/Qualifiers  
1. .963  
/organism="Mus musculus"  
/strain="BFM/2Msf"  
/sub-species="brevirostris"  
/db\_xref="taxon:10090"  
/chromosome="7"  
/note="sequence used for primer design based on  
Acc#AF064792"  
1. .963  
/gene="FUT2"  
<1. .>963  
/gene="FUT2"  
/EC\_number="2.4.1.69"  
/note="Sec2: Se"  
/codon\_start=1  
/product="GDP-L-fucose:beta-D-galactoside  
2-alpha-L-fucosyltransferase"  
/protein\_id="BAB68639.1"  
/db\_xref="GI:15822975"  
/translation="FLIFVVTIIHLOQRIVKLTLSKELQAVOMSPNARTDM  
QOSAKLOGIFTINSIGLNCMGVATLALARMNGRLAPFESMHNALAPFISLP  
VLHSDTARRIPWGNHLDNWEERYRIPQYVRFTGPCSWTFYHLRPELKEFTL  
HDVREQAFLRWLRVNGSQPSIFVGVHVRGRDYGVMKPKVGVVDRGLYLEKALD  
RFRARYSPFVYVTSNGMKNRENTSLGDDVFNAGNIEGSPAKDFALLTQCNTHTM  
TIGTFGTAAYLAGDITLYLANITLPSPLKIFKPAALPEWM"  
214 a 285 c 257 g 207 t

BASE COUNT  
ORIGIN  
Query Match 71.6%; Score 764.6; DB 10; Length 963;  
Best Local Similarity 89.9%; Pred. No. 2.2e-190;

Matches 850; Conservative 0; Mismatches 74; Indels 2; Gaps 2;  
QY 1 CTCACAGCGCAATAGTGAAGCTCCAAACCCCTGTTCAGAGAAGGAATTAACGATGACGACT 60  
Db 40 CTCACAGCAACGAATAGTGAAGCTCCAAACCCCTGTTCAGAGAAGGAATTAAC--AGCGGTT 96  
QY 61 CAAATGTCTCTGGGAAACACAGAAACCCAGAGATGCGACGGGACAGCGAGCAGCATGGG 120  
Db 97 CAAATGTCTCTCACCAACGCGGCAAGAACAGACATGCGACGA-----G 138  
QY 121 AATGAGAGCTGCGGGGCAATGTTACGATCAATTCATTTGGCGGCTGGGAACAGAGATG 180  
Db 139 AGTGCAAGCTGCAGGGCATATTCAGATCAATTCATTCGCGCGCTGGGNAACAGATG 198  
QY 181 GCGAATACGCAACACTCTTTTGCATCTGCGAGGATGAACGGACGGCTTGGCTTCATCCGC 240  
Db 199 GCGAATATGCTACATGTTTTCATGCGCCAGGATGAACGGCTTGGCTTCCCTTCATCCCT 258  
QY 241 GCATCCATGCACAACGCTCTAGCGCCCATCTTCAGGATCAGCTCCCGTGTTCACACAGC 300  
Db 259 GAATCCATGCACAGCTCTAGCGCCCATCTTCAGGATCAGCTCCCGTGTTCACACAGC 318  
QY 301 GACAGCGCAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGAGCGT 360  
Db 319 GACAGCGCAAGAGATCCCGTGGCAGAAATACCATCTCAACGACTGGATGGAGAGCGT 378  
QY 361 TACCGCCACATTCGGGACACTTTTGGCTTCAGGGATACCCCTGCTTCGACCTTC 420  
Db 379 TACCGCCACATTCGGGACACTTTTGGCTTCAGGGATACCCCTGCTTCGACCTTC 438  
QY 421 TACCACCACTTCGGCCAGAGATCCCTGAAGAGTTCACCTGATGACCAACGCTGGGAG 480  
Db 439 TACCACCACTTCGGCCAGAGATCCCTGAAGAGTTCACCTGATGACCAACGCTGGGAG 498  
QY 481 GAGCCCAAGGCTTCTCTCGTGGTGTGCGGGTGAATGGAGCCAGGAGTCTTTGTG 540  
Db 499 GAGCCCAAGGCTTCTCTCGTGGTGTGCGGGTGAATGGAGCCAGGAGTCTTTGTG 558  
QY 541 GGTGTTCATGTGCGCGAGGAGTATGTGATGTCATGCTGCTGCTGCTGCTGCTGCTGCTG 600  
Db 559 GGTGTTCATGTGCGCGAGGAGTATGTGATGTCATGCTGCTGCTGCTGCTGCTGCTGCTG 618  
QY 601 GTGGTTCACCGGGTTCCTGGAAGAGCCCTGATGTTCCGGGACGCTATCACT 660  
Db 619 GTGGTTCACCGGGTTCCTGGAAGAGCCCTGATGTTCCGGGACGCTATCACT 678  
QY 661 CAGTCTTCGTGTTTACAAGCAACGGTATGGCTTGGTGGGAGAGCAATTAATGCTTCC 720  
Db 679 CAGTCTTCGTGTTTACAAGCAACGGTATGGCTTGGTGGGAGAGCAATTAATGCTTCC 738  
QY 721 CGAGGAGACGTGGTGTTCGCGGCAATGGTATTGAGGGGTCCGACGCAAGGACTTCGCG 780  
Db 739 CTAGGAGACGTGGTGTTCGCGGCAATGGTATTGAGGGGTCCGACGCAAGGACTTCGCG 798  
QY 781 CTGCTCACCAGTGCACACCAACCATCATGACTATTGGGACTTTGGGATTTGGGCTGCC 840  
Db 799 CTCTCACCAGTGCACACCAACCATCATGACTATTGGGACTTTGGGATTTGGGCTGCC 858  
QY 841 TACCTGGCAGTGGTGATACCATCTACTTACGCAACTACACCTTCCGGATTTCCTCGTTC 900  
Db 859 TACCTGGCAGTGGTGATACCATCTACTTACGCAACTACACCTTCCGGATTTCCTCGTTC 918  
QY 901 CTCGAAGTCTTTAAGCCAGAGGACGCTTCTCCTACCCGAATGGGTG 945  
Db 919 CTCGAAGTCTTTAAGCCAGAGGACGCTTCTCCTCCTCCGAGTGGATG 963

RESULT 11  
AB039116  
LOCUS  
DEFINITION  
ACCESSION

AB039116 963 bp DNA linear ROD 02-OCT-2001  
Mus musculus Fut2 gene for Gdp-L-fucose:beta-D-galactoside  
2-alpha-L-fucosyltransferase, partial cds, strain:BLG2/Msf.  
AB039116

```

AB039116.1 GI:15822977
FUT2; GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase.
Mus musculus (sub_species:musculus, strain:BLG2/Msf) DNA.
Mus musculus
ORGANISM
REFERENCE
AUTHORS
Liu,Y., Kitano,T., Koide,T., Shiroishi,T., Moriaki,K. and
Saitou,N.
Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes
of Five Mus musculus subspecies
Unpublished
JOURNAL
REFERENCE
AUTHORS
Liu,Y. and Saitou,N.
Direct Submission
JOURNAL
Submitted (29-FEB-2000) Naruya Saitou, National Institute of
Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
URL:http://sayer.lab.nig.ac.jp/, Tel:81-559-81-6790,
Fax:81-559-81-6789)
FEATURES
Location/Qualifiers
1..963
/organism="Mus musculus"
/strain="BLG2/Msf"
/sub_species="musculus"
/db_xref="taxon:10090"
/chromosome="7"
/note="sequence used for primer design based on
Acc#AF064792"
1..963
/gene="FUT2"
<1..>963
/gene="FUT2"
/EC_number="2.4.1.69"
/note="Sec2; Se"
/codon_start=1
/product="GDP-L-fucose:beta-D-galactoside
2-alpha-L-fucosyltransferase"
/protein_id="BAB68640.1"
/db_xref="GI:15822978"
/translation="ELIFVFIYTIHLOORIVKLTSEKLEQAVOMSSPNAFTDM
QQSALQFTINSIGRLNGQEVATLFAALMNGRLAFIPESMENALAFIRSLP
VLHSTARIQPNVHLNDWBEERHPIQVRYETGPCSWTFYHLRPEILKEFTL
HDHVRPEAQRLGLRVNGSQSTFVGVVRGDDYVHVMPKVGKVVADRGYLEKALD
REFARYSPFVVTSGMVCWRENTSLGVDVFVAGNGEGSPAKDFALLTQCNHTIM
TIGTGIAWYLAGDITVILANYTLPSDFLIFKFAALFPEWM"
BASE COUNT 214 a 285 c 257 g 207 t
ORIGIN
Query Match 71.6%; Score 764.6; DB 10; Length 963;
Best Local Similarity 89.9%; Pred. No. 2.2e-190;
Matches 850; Conservative. 0; Mismatches 74; Indels 21; Gaps 2;
QY 1 CTCACACAGGATAGTGAGCTCCAAACCCCTGTCAGAGGAAGGANTACGATGACGAT 60
Db 40 CTCACACAGGATAGTGAGCTCCAAACCCCTGTCAGAGGAAGGANTAC---AGCGGGT 96
QY 61 CAAATGCTCTCGGGAACACAGAAAGCCAGAGATCGGACGCGGACAGCAGCATGGG 120
Db 97 CAAATGCTCTCCAAACCGGCGGAGAGACAGATGACGCA-----G 138
QY 121 AATGGAGAGCTGGGGGCGATGTTCCAGATCAATTCATTGGCGGTGGGGAACACGATG 180
Db 139 AGTGCCAAGCTGAGGGCATATTCAGATCAATTCATCGCGGCTGGGGAACACGATG 198
QY 181 GCGGATAGCCACATCTTTGCACTGGCAGATGAACGGACGGCTTGGTTTCATCCCC 240
Db 199 GCGGAATATGCTACATGTTTGCACTGGCAGATGAACGGTGGCTTGCCTTCATCCCT 258
QY 241 GCATCCATGACAAACGCTCTAGCGCCATCTTCAGGATCAGCTCCCGGTGTTACACAG 300
Db 259 GAATCCATGACAAACGCTCTAGCGCCATCTTCAGGATCAGCTCCCGGTGTTACACAG 318

```

```

QY 301 GACACGGCCAAAAGATCCATGGCAGAAATACCATCTCAACGACTGGAGGAGCGT 360
Db 319 GACACAGCAGAGAGATCCGTGGCAGAAATACCATCTCAACGACTGGAGGAGCGT 378
QY 361 TACCGCCACATTCGCGGACACATTTGTGCGCTTCACGGGATACCCGTGCTCTCGACCTTC 420
Db 379 TACCGCCACATTCGCGGCGAGTATGTGCGTTTACCGGATACCCGTGCTCTCGACCTTC 438
QY 421 TACCACCACTCGGCCCCAGAGATCCTCAAGGAGTTCACCTGCATCACCACGCTGGGGAG 480
Db 439 TACCACCACTCGGCCCCAGAGATCCTCAAGGAGTTCACCTGCATCACCACGCTGGGGAG 498
QY 481 GAGGCCACAGCGCTTCCTGCGTGGTCTGCGGGTGAATGGGAGCCAGCCGAGTACTTTGTG 540
Db 499 GAGGCCACAGCTTCCTGCGTGGTCTGCGGGTGAATGGGAGCCAGCCGAGTACTTTGTG 558
QY 541 GGTGTCATGTGCGCGCGAGGGGACTATGTGCATGTATGCTCATGCTTAAATGTGGAAGGCGTG 600
Db 559 GGTGTCATGTGCGCGCGAGGGGACTATGTGCATGTATGCTCATGCTTAAATGTGGAAGGCGTG 618
QY 601 GTGGCTGACCGGGGTTACCTGGAAAGGCCCTGGATATGTCGGGCGACGCTATTCATCT 660
Db 619 GTGGCTGACCGGGGTTACCTGGAAAGGCCCTGGAGAGTTCGGGCGACGCTATTCATCT 678
QY 661 CCAGTCTTCGTGTTTACAAGCAACGGTATGGCTGGTGGCGGAGAACATTAAGTCTCC 720
Db 679 CCAGTCTTCGTGTTTACAAGCAACGGTATGGCTGGTGGCGGAGAACATTAAGTCTCC 738
QY 721 CGAGGAGAGCTGTGTTTCGCGGCAATGTTATGAGGGTGGCGGACCAAGGACTTCGGG 780
Db 739 CTAGGAGAGCTGTGTTTCGCGGCAATGTTATGAGGGTGGCGGACCAAGGACTTCGGG 798
QY 781 CTGCTCACCAGTGCACACACCATCATGACTATTTGGGACCTTTGGGATTTGGGTGGC 840
Db 799 CTCTCACCAGTGCACACACCATCATGACTATTTGGGACCTTTGGGATTTGGGTGGC 858
QY 841 TACCTGGCAGGTGGTGATACCATCTACTTAGCCAACTACACCTTCGCGATTCGCTTC 900
Db 859 TACCTGGCAGGTGGTGATACCATCTACTTAGCCAACTACACCTTCGCGATTCGCTTC 918
QY 901 CTCAAAGCTTTAAGCCAGGAGCGAGCTTCCTTACCACGGAATGGTG 945
Db 919 CTCAAATCTTTAAGCCAGGAGCGAGCTTCCTTACCACGGAATGGTG 963
RESULT 12
AB039117
LOCUS
DEFINITION
Mus musculus FUT2 gene for GDP-L-fucose:beta-D-galactoside
2-alpha-L-fucosyltransferase, partial cds, strain:CAST/Ei.
AB039117
ACCESSION
AB039117.1 GI:15822979
VERSION
FUT2; GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase.
Mus musculus (sub_species:castaneus, strain:CAST/Ei) DNA.
KEYWORDS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ORGANISM
REFERENCE
1
AUTHORS
Liu,Y., Kitano,T., Koide,T., Shiroishi,T., Moriaki,K. and
Saitou,N.
Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes
of Five Mus musculus subspecies
Unpublished
JOURNAL
REFERENCE
2 (bases 1 to 963)
AUTHORS
Liu,Y. and Saitou,N.
Direct Submission
JOURNAL
Submitted (29-FEB-2000) Naruya Saitou, National Institute of
Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
URL:http://sayer.lab.nig.ac.jp/, Tel:81-559-81-6790,
Fax:81-559-81-6789)
FEATURES
Location/Qualifiers
1..963
source

```



```

RESULT 14
AB039119 214 a 285 c 257 g 207 t
LOCUS      963 bp      DNA      linear      ROD 02-OCT-2001
DEFINITION Mus musculus FUT2 gene for GDP-L-fucose:beta-D-galactoside
            2-alpha-L-fucosyltransferase, partial cds, strain:pgn2.
ACCESSION  AB039119.1 GI:15822983
VERSION     AB039119.1
KEYWORDS    FUT2; GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase.
SOURCE      Mus musculus (sub_species:domesticus, strain:pgn2) DNA.
ORGANISM    Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
AUTHORS     Liu,Y., Kitano,T., Koide,T., Shiroishi,T., Moriaki,K. and
            Saitou,N.
TITLE       Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes
            of Five Mus musculus subspecies
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 963)
AUTHORS     Liu,Y. and Saitou,N.
TITLE       Direct Submission
JOURNAL     Submitted (29-FEB-2000) Naruya Saitou, National Institute of
            Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,
            Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
            URL:http://sayer.lab.nig.ac.jp/, Tel:81-559-81-6790,
            Fax:81-559-81-6789)
FEATURES             Location/Qualifiers
     source           1..963
                     /organism="Mus musculus"
                     /strain="pgn2"
                     /sub_species="domesticus"
                     /db_xref="taxon:10090"
                     /chromosome="7"
     note="sequence used for primer design based on
           Acc#AF064792"
     gene             1..963
     CDS               1..963
                     /gene="FUT2"
                     /EC_number="2.4.1.69"
                     /note="Sec2; Se"
                     /codon_start=1
                     /product="GDP-L-fucose:beta-D-galactoside
                     2-alpha-L-fucosyltransferase"
                     /protein_id="BAB58643.1"
                     /db_xref="GI:15822984"
                     /translation="FLIEFVFTSTIIHQRIKVLQTLSEKELQAVQSSNPAATDM
                     QCSAKLOGIFTINSGRLGNQMEYATILFALAMNRLAFIPESMHALAIPRIISLP
                     VLHSTARRIPWNYHLNDWMEERYHPIGOYVRFYPCSWTFYHHLRPEILFETL
                     HDHVEEAQAPLRGLRVNGSOPSTFVGVRGDDYHVMKPVKGVADRGYLEKALD
                     RFRARYSSPVFTVTSNGMAGCFENITSLGSDVFPAGNIEGSPAKDFALLTQCNHTIM
                     TIGTGIWAAYLAGDITILANLYTLPDPSFLKIFKPAALFLENH"
BASE COUNT  214 a 285 c 257 g 207 t
ORIGIN
Query Match 71.6%; Score 764.6; DB 10; Length 963;
Best Local Similarity 89.9%; Pred. No. 2.2e-190;
Matches 850; Conservative 0; Mismatches 74; Indels 21; Gaps 2;
QY 1 CTCACGACGCGAATAGTGAAGCTCCAAACCCCTGTCCAGAGAGGAATATCCGATGACGACT 60
Db 1 CTCACGACGCGAATAGTGAAGCTCCAAACCCCTGTCCAGAGAGGAATATCCGATGACGACT 60
QY 40 CTCACGACGCGAATAGTGAAGCTCCAAACCCCTGTCCAGAGAGGAATATCCGATGACGACT 96
Db 40 CTCACGACGCGAATAGTGAAGCTCCAAACCCCTGTCCAGAGAGGAATATCCGATGACGACT 96
QY 61 CAATGTCTCTCGGGAACACAGAACCCCGAGATGCGGACGACGACGACGACGATGG 120
Db 61 CAATGTCTCTCGGGAACACAGAACCCCGAGATGCGGACGACGACGACGACGATGG 120
QY 97 CAATGTCTCTCGGGAACACAGAACCCCGAGATGCGGACGACGACGACGATGG 138
Db 97 CAATGTCTCTCGGGAACACAGAACCCCGAGATGCGGACGACGACGACGATGG 138
QY 121 AATGAGAGCTCGGCGCATGTTCACGATCAATTCATTTGGCGGCTGGGAGACGATG 180
Db 121 AATGAGAGCTCGGCGCATGTTCACGATCAATTCATTTGGCGGCTGGGAGACGATG 180
QY 139 AGTGCCACGCTGACGGGCATATTCAGATCAATTCATTCGCGGCTGGGAGACGATG 198
Db 139 AGTGCCACGCTGACGGGCATATTCAGATCAATTCATTCGCGGCTGGGAGACGATG 198
QY 181 GCGAATAGCCACACTCTTTGCACCTGGCAGGATGACGAGGAGGCTTGGCTTCATCC 240
Db 181 GCGAATAGCCACACTCTTTGCACCTGGCAGGATGACGAGGAGGCTTGGCTTCATCC 240

```

```

RFRARYSSPVFTVTSNGMAGCFENITSLGSDVFPAGNIEGSPAKDFALLTQCNHTIM
TIGTGIWAAYLAGDITILANLYTLPDPSFLKIFKPAALFLENH"
BASE COUNT  214 a 285 c 257 g 207 t
ORIGIN
Query Match 71.6%; Score 764.6; DB 10; Length 963;
Best Local Similarity 89.9%; Pred. No. 2.2e-190;
Matches 850; Conservative 0; Mismatches 74; Indels 21; Gaps 2;
QY 1 CTCACGACGCGAATAGTGAAGCTCCAAACCCCTGTCCAGAGAGGAATATCCGATGACGACT 60
Db 1 CTCACGACGCGAATAGTGAAGCTCCAAACCCCTGTCCAGAGAGGAATATCCGATGACGACT 60
QY 40 CTCACGACGCGAATAGTGAAGCTCCAAACCCCTGTCCAGAGAGGAATATCCGATGACGACT 96
Db 40 CTCACGACGCGAATAGTGAAGCTCCAAACCCCTGTCCAGAGAGGAATATCCGATGACGACT 96
QY 61 CAATGTCTCTCGGGAACACAGAACCCCGAGATGCGGACGACGACGACGACGATGG 120
Db 61 CAATGTCTCTCGGGAACACAGAACCCCGAGATGCGGACGACGACGACGACGATGG 120
QY 97 CAATGTCTCTCGGGAACACAGAACCCCGAGATGCGGACGACGACGACGATGG 138
Db 97 CAATGTCTCTCGGGAACACAGAACCCCGAGATGCGGACGACGACGACGATGG 138
QY 121 AATGAGAGCTCGGCGCATGTTCACGATCAATTCATTTGGCGGCTGGGAGACGATG 180
Db 121 AATGAGAGCTCGGCGCATGTTCACGATCAATTCATTTGGCGGCTGGGAGACGATG 180
QY 139 AGTGCCACGCTGACGGGCATATTCAGATCAATTCATTCGCGGCTGGGAGACGATG 198
Db 139 AGTGCCACGCTGACGGGCATATTCAGATCAATTCATTCGCGGCTGGGAGACGATG 198
QY 181 GCGAATAGCCACACTCTTTGCACCTGGCAGGATGACGAGGAGGCTTGGCTTCATCC 240
Db 181 GCGAATAGCCACACTCTTTGCACCTGGCAGGATGACGAGGAGGCTTGGCTTCATCC 240

```



```
199 GCGAATATGCTACATTTGCTACCTGGCCAGGATGACGGTGGCTTCATCCCT 258
241 GCATCCATGACACAGCTCTAGCGCCCATCTTTCAGATCAGCTCCCGGTGTTACACAGC 300
259 GAATCCATGACACAGCTCTAGCGCCCATCTTTCAGATCAGCTCCCGGTGTTACACAGC 318
301 GACAGGCCAAAGATGCCATGCGAGATTAACATCTCAACGACTGGATGAGAGCGT 360
319 GACACAGCAGAAGATCCCGTGGCGAATTAACACCTCAACGACTGGATGAGAGCGT 378
351 TACGCCACATTCGCGGACACTTTTGTGGCTTCACGGGATACCGGTGCTCTGGACCTTC 420
379 TACGCCACATTCGCGGACACTTTTGTGGCTTCACGGGATACCGGTGCTCTGGACCTTC 438
421 TACACACACCTTGCGCCAGAGATCTGAAGAGTTCACCTGATGACCACTGGCGGAG 480
439 TACCACACCTTGCGCCAGAGATCTGAAGAGTTCACCTGATGACCACTGGCGGAG 498
481 GAGGCCAGGCTTCCTGCGTGGTCTGCGGTGATGGAGCCAGCCGAGTCTTTGTG 540
499 GAGGCCAGGCTTCCTGCGTGGTCTGCGGTGATGGAGCCAGCCGAGTCTTTGTG 558
541 GGTGTCCATGTGCGCCAGGAGTCTGATGATGCTGATGCTGATGCTGATGCTGATGCTG 600
559 GGTGTCCATGTGCGCCAGGAGTCTGATGATGCTGATGCTGATGCTGATGCTGATGCTG 618
601 GTGGTGTACCGGGTACCTGGAAAGCCCTGATATGTTCCGGGACAGCTATTCATCT 660
619 GTGGTGTACCGGGTACCTGGAAAGCCCTGATATGTTCCGGGACAGCTATTCATCT 678
661 CCAGTCTTCGTTGTACAGCAACGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
679 CCAGTCTTCGTTGTACAGCAACGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 738
721 CGAGGAGAGCTGGTGTTCGCGGGCAATGATGATGAGGGTTCGCGGAGGACTTCGCG 780
739 CTAGGAGAGCTGGTGTTCGCGGGCAATGATGATGAGGGTTCGCGGAGGACTTCGCG 798
781 CTGCTCACCAGTGAACCAACACATCATGATATGAGGAGTTCGCGGAGGACTTCGCG 840
799 CTGCTCACCAGTGAACCAACACATCATGATATGAGGAGTTCGCGGAGGACTTCGCG 858
841 TACCTGGAGTGGTGTATACCATCTACTAGTACCACTACACCTTCGCGGATTCGCGTTC 900
859 TACCTGGAGTGGTGTATACCATCTACTAGTACCACTACACCTTCGCGGATTCGCGTTC 918
901 CTCAAAGTCTTAAAGCCAGAGGAGCTTCCTTACCCGAATGGTG 945
919 CTCAAAGTCTTAAAGCCAGAGGAGCTTCCTTACCCGAATGGTG 963

RESULT 15
AB039120
LOCUS
DEFINITION
AB039120
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
```

```
Submitted (29-FEB-2000) Naruya Saitou, National Institute of
Genetics, Laboratory of Evolutionary Genetics, 1111 Yata, Mishima,
Shizuoka 411-8540, Japan (E-mail:nsaitou@genetics.nig.ac.jp,
URL:http://sayer.lab.nig.ac.jp/, Tel:81-559-81-6730,
Fax:81-559-81-6789)
Location/Qualifiers
1. 963
/organism="Mus musculus"
/strain="SWN/Msf"
/sub_species="molossinus"
/db_xref="taxon:10090"
/chromosome="7"
/note="sequence used for primer design based on
Acc#AF064792"
1. 963
/gene="PUT2"
<1. >963
/gene="PUT2"
/EC_number="2.4.1.69"
/note="Sec2: Se"
/codon_start=1
/product="GDP-L-fucose:beta-D-galactoside
2-alpha-L-fucosyltransferase"
/protein_id="BAB68644.1"
/db_xref="GI:15822986"
/translation="FLIPVFVITIIHQQRIVKLOTLSKELQAVMSSPNAARDIM
QQSKLOGIFTINSIGELGNOMGEYATILFALRMNGLATIPESMHNALAFIRISLP
VLSDTARRIFWQNYHLNDWERYRHIPGGYVRETYPCSWTYPHHLRPILEFTL
HDVREAFQALRGLRVNQSPTFVGVHVRGDIYVHVMPKVGKVDALLTQCNHTIM
RFRYSFVFTVSNMNCWRENTSLGDFVAGNIGPKSPKADFDALTTQCNHTIM
TIGTFGIWAAYLAGGDTIYLANITLPSFLKIFKPAALFPEWM"
BASE COUNT 214 a 285 c 257 g 207 t
ORIGIN
Query Match 71.6%; Score 764.6; DB 10; Length 963;
Best Local Similarity 89.9%; Pred. No. 2.2e-190;
Matches 850; Conservative 0; Mismatches 74; Indels 21; Gaps 2;
QY 1 CTCGAGCAGCGAATAGTGAAGCTCCAAACCCCTGCAGAGAGGAATTACCGATGACGACT 60
DB 40 CTCGAGCAGCGAATAGTGAAGCTCCAAACCCCTGCAGAGAGGAATTACCGATGACGACT 96
QY 61 CAAATGTCTCTGGGAAACACAGAAAGCCGACGAGATCCGACGGGACAGCAGCAGTGG 120
DB 97 CAAATGTCTCTCACCACACGCGGCAAGACAGACATGACGCA-----G 138
QY 121 ATGAGAGAGCTCGGGGATGTTACAGATCAATTCATTTGGCCGCTGGGGACAGATG 180
DB 139 AGTGCCAAAGCTGCAGGCGATATTCAGATCAATTCATTCGCGGCTGGGGACAGATG 198
QY 181 GCGGAATAGCCACACTCTTTGCACTGGCCAGGATGAACGAGCGGCTTCGCTTCATCCC 240
DB 199 GCGGAATAGTACATTTGTTGCACTGGCCAGGATGAACGAGCGGCTTCGCTTCATCCC 258
QY 241 GCATCCATGACACAGCTCTAGCGCCCATCTTTCAGATCAGCTCCCGGTGTTACACAGC 300
DB 259 GAATCCATGACACAGCTCTAGCGCCCATCTTTCAGATCAGCTCCCGGTGTTACACAGC 318
QY 301 GACAGGCCAAAGATCCCATGCGAGATTAACATCTCAACGACTGGATGAGAGCGT 360
DB 319 GACAGGCCAGAGGATCCCGTGGCGAATTAACACCTCAACGACTGGATGAGAGCGT 378
QY 361 TACCGCCACATTCGCGGACACTTTTGTGGCTTCACGGGATACCGGTGCTCTGGACCTTC 420
DB 379 TACCGCCACATTCGCGGACACTTTTGTGGCTTCACGGGATACCGGTGCTCTGGACCTTC 438
QY 421 TACCACACCTTGCGCCAGAGATCTGAAGAGTTCACCTGATGACCACTGGCGGAG 480
DB 439 TACCACACCTTGCGCCAGAGATCTGAAGAGTTCACCTGATGACCACTGGCGGAG 498
QY 481 GAGGCCAGGCTTCCTGCGTGGTCTGCGGTGATGGAGCCAGCCGAGTCTTTGTG 540
DB 499 GAGGCCAGGCTTCCTGCGTGGTCTGCGGTGATGGAGCCAGCCGAGTCTTTGTG 558
```

JOURNAL  
source  
FEATURES

1. 963  
/organism="Mus musculus"  
/strain="SWN/Msf"  
/sub\_species="molossinus"  
/db\_xref="taxon:10090"  
/chromosome="7"  
/note="sequence used for primer design based on  
Acc#AF064792"  
1. 963  
/gene="PUT2"  
<1. >963  
/gene="PUT2"  
/EC\_number="2.4.1.69"  
/note="Sec2: Se"  
/codon\_start=1  
/product="GDP-L-fucose:beta-D-galactoside  
2-alpha-L-fucosyltransferase"  
/protein\_id="BAB68644.1"  
/db\_xref="GI:15822986"  
/translation="FLIPVFVITIIHQQRIVKLOTLSKELQAVMSSPNAARDIM  
QQSKLOGIFTINSIGELGNOMGEYATILFALRMNGLATIPESMHNALAFIRISLP  
VLSDTARRIFWQNYHLNDWERYRHIPGGYVRETYPCSWTYPHHLRPILEFTL  
HDVREAFQALRGLRVNQSPTFVGVHVRGDIYVHVMPKVGKVDALLTQCNHTIM  
RFRYSFVFTVSNMNCWRENTSLGDFVAGNIGPKSPKADFDALTTQCNHTIM  
TIGTFGIWAAYLAGGDTIYLANITLPSFLKIFKPAALFPEWM"

QY 541 GGTGTCATGTGCGCCGAGGGGACTATGTGTCATGTGCTATGCTAATGTGIGGAGGGGCTG 600  
Db |||||  
QY 559 GGTGTCATGTGCGCCGAGGGGACTATGTGTCATGTGCTAATGTGIGGAGGGGCTG 618  
Db |||||  
QY 601 GTGGCTGACCGGGGTTACCTGGAAAGGCCCTGGATATGTTCGGGGACGCTATTCTCT 660  
Db |||||  
QY 619 GTGGCTGACCGGGGTTACCTAGNAAAGGCCCTGGACAGGTTCCGGGCACGCTATTCTCT 678  
Db |||||  
QY 661 CCACTCTTCGTGGTTACAAGCAACGGTATGGCCCTGGTCCGGGAGACATTAATGCTTCC 720  
Db |||||  
QY 679 CCACTCTTCGTGGTTACAAGCAACGGTATGGCCCTGGTCCGGGAGACATTAACACCTCC 738  
Db |||||  
QY 721 CGAGGAGACGTGGTGTGCGGGGCAATGTTATGAGGGGTGCGCCAGCCAGGACTTCCG 780  
Db |||||  
QY 739 CTAGGAGACGTGGTGTGCGGGGCAATGTTATGAGGGGTGCGCCAGCCAGGACTTCCG 798  
Db |||||  
QY 781 CTGCTCACCCAGTGCACACCACCATCATGACTATTTGGGACCTTTGGGATTTGGGCTGCC 840  
Db |||||  
QY 799 CTCCTCACCCAGTGCACACCACCATCATGACTATTTGGGACCTTTGGGATTTGGGCTGCC 858  
Db |||||  
QY 841 TACCTGGCAGTGGTGGTATACCATCTACTAGCCAACTACACCCCTCCGGATTCCTCCGTC 900  
Db |||||  
QY 859 TACCTGGCAGTGGTGGTATACCATCTACTAGCCAACTACACCCCTCCGGATTCCTCCGTC 918  
Db |||||  
QY 901 CTCRAAGTCTTTAAGCCAGAGGCGCCCTTCTTACCCGAATGGTG 945  
Db |||||  
QY 919 CTCRAATCTTTAAGCCAGAGGCGCCCTTCTTACCCGAATGGTG 963  
Db |||||

Search completed: May 26, 2003, 14:49:17  
Job time : 2883.28 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 09:17:27 : Search time 276.996 Seconds  
(without alignments)  
8882.922 Million cell updates/sec

Title: US-10-040-863-9  
Perfect score: 1068  
Sequence: 1 ctcacagcagcaatagttaa.....gaagagccttctgatggagg 1068

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002:\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1068	100.0	1068	21 AAC67966	Rat hepatoma H35 c
2	1068	100.0	1068	24 AAD27208	Rat hepatoma H35 c
3	1068	100.0	1149	21 AAC67965	Rat hepatoma H35 c
4	1068	100.0	1149	24 AAD27207	Rat hepatoma H35 c
5	612.8	57.4	1043	19 AAV29003	Porcine secretor t
6	612.8	57.4	1043	19 AAV21639	Pig secretor encod
7	590.6	55.3	2115	19 AAV38323	Human Sec2 coding
8	589.4	55.2	3088	24 ABL64735	Lung cancer relate
9	589.4	55.2	3088	24 ABL66311	Lung cancer relate

10	589.4	55.2	3088	24	ABL70026	Pancreas cancer re
11	423.8	39.7	1098	19	AAV21640	Pig H transferase
12	423.8	39.7	1269	20	AAV15872	Fragment of pH83x
13	423.8	39.7	1269	20	AAV03811	Swine alpha-1,2-fu
14	423.8	39.7	1269	21	AAZ94417	Pig alpha-1,2 fuc
15	423.8	39.7	1269	24	AAI72831	FU1 cDNA. Sus sc
16	405	37.9	1155	16	AAQ98461	GDP-L-fucose-beta-
17	405	37.9	1155	16	AAQ01082	2-Alpha-fucosyltra
18	405	37.9	1155	16	AAQ01083	Human H-transferas
19	405	37.9	1174	17	AAI12238	Human H-transferas
20	405	37.9	1174	21	AAV53820	Mouse H2kb gene pr
21	405	37.9	2268	18	AAV63576	Human alpha 1,2 fu
22	405	37.9	3373	18	AAV67972	Chicken beta-actin
23	405	37.9	3791	18	AAV63575	GDP-Fuc:beta-D-gal
24	405	37.9	8174	12	AAQ13332	Human alpha(1,2)-f
25	405	37.9	8174	18	AAV61677	DNA encoding a g-y
26	403.4	37.8	8174	15	AAQ56908	X. laevis alpha1,
27	304.8	28.5	1071	20	AAV7268	Human secreted pro
28	125.4	11.7	473	21	AAV03536	Rabbit alpha1-2fuc
29	72.8	6.8	100	21	AAV67984	Rabbit alpha1-2fuc
30	71.2	6.7	100	21	AAV67978	Rabbit alpha1-2fuc
31	71.2	6.7	100	21	AAV67981	Rabbit alpha1-2fuc
32	69.6	6.5	100	21	AAV67975	Human alpha1-2fuc
33	63.2	5.9	100	21	AAV67969	Human alpha1-2fuc
34	63.2	5.9	100	21	AAV67972	Human alpha1-2fuc
35	44	4.1	495	22	AAH06199	Human cDNA clone (
36	42.8	4.0	4069	22	AAH17772	Human cDNA sequenc
37	42.8	4.0	5236	22	AAH16566	Human polynucleoti
38	42.8	4.0	5530	22	AAK52640	Human polynucleoti
39	40.4	3.8	2297	22	AAV05087	Human secreted pro
40	40.4	3.8	2612	22	AAV05059	Human secreted pro
41	40.4	3.8	2953	23	ABK43445	DNA encoding novel
42	40.4	3.8	3044	22	ABK31284	Human cDNA encodin
43	40.4	3.8	3044	24	ABO66608	Human polynucleoti
44	38.4	3.6	50	21	AAV67971	Human alpha1-2fuc
45	38.2	3.6	51	22	AAV79569	Human DNA containi

ALIGNMENTS

RESULT 1  
AAC67966  
ID AAC67966 standard; cDNA; 1068 BP.

XX AAC67966;

XX 19-FEB-2001 (first entry)

XX Rat hepatoma H35 cell alpha1-2fucosyltransferase catalytic domain cDNA.

XX Rat; alpha1-2fucosyltransferase; cytosolic; neuroprotective;

XX nontropic; gene therapy; Fucalalpha1-2Galbeta1-3GalNAc; immunotherapy;

XX immunosuppression; cancer; neurological disease;

XX small cell lung carcinoma; ss.

XX Rattus norvegicus.

XX WO2000064464-A1.

XX 02-NOV-2000.

XX 23-APR-1999; 99WO-US07384.

XX 23-APR-1999; 99WO-US07384.

XX (PACI-) PACIFIC NORTHWEST CANCER FOUND.

XX Holmes EH, Sherwood AL;

XX WPI: 2000-687262/67.

XX P-PSDB; AAB36105.

PT New rat ganglioside GM1-specific alpha1-2Fucosyltransferase, useful for  
PT preparation of fucosyl GM1 which is useful as a nutritional composition  
XX or immunotherapeutic for cancer and neurological diseases -  
XX Claim 11; Fig 3A; 91pp; English.

XX The present sequence is given in a specification relating to a rat  
CC ganglioside GM1-specific alpha1-2fucosyltransferase protein. The protein  
CC or its cellular fraction is useful for synthesis of a molecule comprising  
CC Fucalphan-2galactose-3galactose, a glycolipid, glycoprotein, glycolipoprotein  
CC or a free oligosaccharide comprising Fucalphan-2galactose-3galactose.  
CC The method involves contacting alpha1-2fucosyltransferase with GDP-fucose  
CC and a molecule or glycolipid, glycoprotein, glycolipoprotein or  
CC oligosaccharide having a terminal Galbeta1-3GalNAc group. It is also  
CC useful for synthesis of fucosyl-GM1 by contacting the protein with  
CC GDP-fucose and ganglioside GM1. The obtained glycoproteins,  
CC glycolipoproteins, glycolipids and oligosaccharides are useful as  
CC nutritional compositions and fucosyl-GM1 is useful for inducing an  
CC immunotherapeutic or immunosuppressive action against cancer,  
CC neurological disease or small cell lung carcinoma.

SQ Sequence 1068 BP; 237 A; 308 C; 300 G; 223 T; 0 other;

Query Match 100.0%; Score 1068; DB 21; Length 1068;

Best Local Similarity 100.0%; Pred. No. 8.5e-289;

Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCACGACGAGTATAGTACGATCTCAACCCCTGTGACGAGAGGAATACCGATGACGACT 60  
DB 1 CTCACGACGAGTATAGTACGATCTCAACCCCTGTGACGAGAGGAATACCGATGACGACT 60  
QY 61 CAATGTCTCTGGGAAACACAGAAAGCCAGAGATGCGACGGGACGAGCAGCAGATGG 120  
DB 61 CAATGTCTCTGGGAAACACAGAAAGCCAGAGATGCGACGGGACGAGCAGCAGATGG 120  
QY 121 AATGAGAGCTCGGGGGATGTTACGATCAATTCATTGGCCGCTGGGGAACACGATG 180  
DB 121 AATGAGAGCTCGGGGGATGTTACGATCAATTCATTGGCCGCTGGGGAACACGATG 180  
QY 181 GCGAATAGCCACACTCTTTCACCTGCGGAGATGAACGGAGCGCTTCCTTCATCC 240  
DB 181 GCGAATAGCCACACTCTTTCACCTGCGGAGATGAACGGAGCGCTTCCTTCATCC 240  
QY 241 GCATCCATGCACACGCTCTACGCGCCATCTCAGATCAGCTCCCGTGTATACACAG 300  
DB 241 GCATCCATGCACACGCTCTACGCGCCATCTCAGATCAGCTCCCGTGTATACACAG 300  
QY 301 GACACGGCCAAAGATCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGCGT 360  
DB 301 GACACGGCCAAAGATCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGCGT 360  
QY 361 TACCGCCACATTCGGGACACTTTGCGCTTCAGGGATACCCGTCCTGGACCTTC 420  
DB 361 TACCGCCACATTCGGGACACTTTGCGCTTCAGGGATACCCGTCCTGGACCTTC 420  
QY 421 TACCGCCACCTGGCCAGAGATCTCTGAAGAGATTCACCTGCATGACCACTGGGGAG 480  
DB 421 TACCGCCACCTGGCCAGAGATCTCTGAAGAGATTCACCTGCATGACCACTGGGGAG 480  
QY 481 GAGGCCAGCCCTTCTCGCTGGTCTGCGGGTGAATGGAGCCAGCCGAGTACTTTTGTG 540  
DB 481 GAGGCCAGCCCTTCTCGCTGGTCTGCGGGTGAATGGAGCCAGCCGAGTACTTTTGTG 540  
QY 541 GGTGTCCATGTGCGGAGGAGTATGTGCATGTCATGCTCCTAATGTGGAGGCGGTG 600  
DB 541 GGTGTCCATGTGCGGAGGAGTATGTGCATGTCATGCTCCTAATGTGGAGGCGGTG 600  
QY 601 GTGGCTACCGGGTACCTCTGAAAGGCGCTGGATATGTCGGGACGCTATTCATCT 660  
DB 601 GTGGCTACCGGGTACCTCTGAAAGGCGCTGGATATGTCGGGACGCTATTCATCT 660  
QY 661 CCAGTCTCTGTTGTACAGCAACGATGGCTGTGCGGGAGAACATTAATGCTTCC 720  
DB 661 CCAGTCTCTGTTGTACAGCAACGATGGCTGTGCGGGAGAACATTAATGCTTCC 720

DB 661 CCAGTCTCTGTTGTACAGCAACGATGGCTGTGCGGGAGAACATTAATGCTTCC 720  
QY 721 CGAGAGACGCTGGTGTTCGCGGGCAATGTTAGGGGTCCCGCAAGACTTCGGG 780  
DB 721 CGAGAGACGCTGGTGTTCGCGGGCAATGTTAGGGGTCCCGCAAGACTTCGGG 780  
QY 781 CTGCTCACCCAGTGCACACACACCATCATGACTATTGGGACCTTTGGGATTTGGCTGC 840  
DB 781 CTGCTCACCCAGTGCACACACACCATCATGACTATTGGGACCTTTGGGATTTGGCTGC 840  
QY 841 TACCTGGCAGGGGTGATACCATCTTACTAGCAACTACACCTTCGCGATTCGCTTC 900  
DB 841 TACCTGGCAGGGGTGATACCATCTTACTAGCAACTACACCTTCGCGATTCGCTTC 900  
QY 901 CTCAAGTCTTTAAGCCAGAGCAGCTTCCTACCCGAATGGTGGGATCCCTGCCGAT 960  
DB 901 CTCAAGTCTTTAAGCCAGAGCAGCTTCCTACCCGAATGGTGGGATCCCTGCCGAT 960  
QY 961 CTCTCCCACTCTTAAAGSCATTAACACACGAGCTGTCTCTGCTCCCACTTCCACCTCAAG 1020  
DB 961 CTCTCCCACTCTTAAAGSCATTAACACACGAGCTGTCTCTGCTCCCACTTCCACCTCAAG 1020  
QY 1021 GCATAAGAGGACTCTGTTTACCTGCGCAGGAGAGCCCTTCTGATGGAA 1068  
DB 1021 GCATAAGAGGACTCTGTTTACCTGCGCAGGAGAGCCCTTCTGATGGAA 1068

RESULT 2

AAD27208

ID AAD27208 standard; DNA; 1068 BP.

XX AAD27208;

XX AC

XX 09-APR-2002 (first entry)

XX DE

XX Rat hepatoma H35 cell alpha1-2FucT catalytic domain encoding DNA.

XX KW

XX Rat; alpha1-2fucosyltransferase; alpha1-2FucT; antisense therapy;

XX KW galactose beta1-3N-acetylgalactosamine; Galbeta1-3GalNAc; glycolipid;

XX KW glycoprotein; glycolipoprotein; oligosaccharide; fucosyl-GM1; cancer;

XX KW gene therapy; oncogenic transformation; cytostatic; ganglioside; GM1;

XX KW cell transformation; catalytic domain; ds.

XX OS

XX Rattus norvegicus.

XX FH

XX Location/Qualifiers

XX CDS

XX 1..1062

XX /tag= a

XX /product= "Rat hepatoma H35 cell alpha1-2FucT catalytic

XX domain"

XX /note= "CDS does not include start codon"

XX FT

XX partial

XX 531..1068

XX /tag= b

XX /note= "Region which overlaps rat FTB"

XX misc\_feature

XX US6329170-B1.

XX PN

XX 11-DEC-2001.

XX XX

XX 23-APR-1999; 99US-0298886.

XX XX

XX 23-APR-1999; 99US-0298886.

XX PR

XX (NWHO-) NORTHWEST HOSPITAL.

XX PA

XX Holmes EH, Sherwood AL;

XX PI

XX WPI; 2002-121132/16.

XX DR

XX P-PSDB; AAE16623.

XX XX

XX Rat hepatoma H35 cell alpha1-2fucosyltransferase, useful for producing

XX GM1-specific alpha1-2fucosyltransferase enzyme by recombinant

XX PT

PT techniques and for detecting oncogenic transformation of test tissues -

XX Claim 2; Fig 3; 41pp; English.

XX The invention relates to rat GM1-specific alphas-2-fucosyltransferase (alpha-2-fucT) enzyme and its corresponding nucleic acid. This nucleic acid is specific for a carbohydrate moiety found in ganglioside GM1, a terminal galactose beta1-3N-acetylgalactosamine (Galbeta1-3GalNAc) saccharide. Alpha-2-fucT DNA is useful for producing rat alpha-2-fucT protein by recombinant techniques. Alpha-2-fucT DNA is useful for the preparative synthesis of fucosyl containing glycolipids, glycoproteins, CC glycolipoproteins and oligosaccharide, and for preparing fucosyl-GM1. CC Alpha-2-fucT DNA is useful for detecting oncogenic transformation which CC involves assaying for changes in expression of alpha-2 fucT. Since CC alpha-2-fucT is activated in cell transformation, antisense sequences CC derived from alpha-2-fucT DNA are useful for inhibiting, suppressing CC or treating cancer. Alpha-2-fucT DNA is useful in gene therapy and CC antisense therapy. The present sequence is rat hepatoma H35 cell CC alpha-2-fucT catalytic domain encoding DNA.

SQ Sequence 1068 BP; 237 A; 308 C; 300 G; 223 T; 0 other;

Query Watch 100.0%; Score 1068; DB 24; Length 1068;

Best Local Similarity 100.0%; Pred. No. 8,6e-289; Mismatches 0; Indels 0; Gaps 0; Matches 1068; Conservative 0;

QY 1 CTCACGACGCAATAGTGAAGCTCCAAACCCCTGTACAGAGAAGTAATACCGATGACGACT 60  
DB 1 CTCACGACGCAATAGTGAAGCTCCAAACCCCTGTACAGAGAAGTAATACCGATGACGACT 60  
QY 61 CAAATGCTCTCGGGAAACACAGAAAGCCAGAGATCGACGGGACAGCAGCATATGGG 120  
DB 61 CAAATGCTCTCGGGAAACACAGAAAGCCAGAGATCGACGGGACAGCAGCATATGGG 120  
QY 121 AATGGAGAGCTCGGGGCACTGTTACAGATCAATTCATTGGCGGCTGGGACACAGATG 180  
DB 121 AATGGAGAGCTCGGGGCACTGTTACAGATCAATTCATTGGCGGCTGGGACACAGATG 180  
QY 181 GCGAATACGCCACACTCTTTGACCTGGCCAGGATGACGGGCTTGGCTTCATCCGC 240  
DB 181 GCGAATACGCCACACTCTTTGACCTGGCCAGGATGACGGGCTTGGCTTCATCCGC 240  
QY 241 GCATCATGCACAAACGCTCTAGCGGCCATCTTACAGATCAGGCTCCGGTGTACACAG 300  
DB 241 GCATCATGCACAAACGCTCTAGCGGCCATCTTACAGATCAGGCTCCGGTGTACACAG 300  
QY 301 GACACGGCCAAAGATCCCATGGCAGATTAACCATCTCAACGACTGGATGGAGGCGT 360  
DB 301 GACACGGCCAAAGATCCCATGGCAGATTAACCATCTCAACGACTGGATGGAGGCGT 360  
QY 361 TACCGCCACATTCGGGACACTTTGTGCGCTTCACGGGATACCGGTCTCCTGGACCTTC 420  
DB 361 TACCGCCACATTCGGGACACTTTGTGCGCTTCACGGGATACCGGTCTCCTGGACCTTC 420  
QY 421 TACACACACTGCGGCCACAGATCTTGAAGAGTTCACCGTGCATGACACAGTGGGGAG 480  
DB 421 TACACACACTGCGGCCACAGATCTTGAAGAGTTCACCGTGCATGACACAGTGGGGAG 480  
QY 481 GAGGCCACGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
DB 481 GAGGCCACGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
QY 541 GGTGTCCATGTGCGGAGGGGACTATGTGCTATGCTATGCTATGCTATGCTATGCTATGCT 600  
DB 541 GGTGTCCATGTGCGGAGGGGACTATGTGCTATGCTATGCTATGCTATGCTATGCTATGCT 600  
QY 601 GTGGCTGACGGGGTTACCTGGAAGAGCCCTGGAATGTTCCGGGACACGCTATTCATCT 660  
DB 601 GTGGCTGACGGGGTTACCTGGAAGAGCCCTGGAATGTTCCGGGACACGCTATTCATCT 660  
QY 661 CCAGTCTCTGCTTACAGACACGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
DB 661 CCAGTCTCTGCTTACAGACACGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720

QY 721 CGAGGAGAGTGTGTTCGGGGCAATGGTATTAGGGGTTCGCCAAGGACTTCGG 780  
DB 721 CGAGGAGAGTGTGTTCGGGGCAATGGTATTAGGGGTTCGCCAAGGACTTCGG 780  
QY 781 CTGCTCACCCAGTGCACACACACCATCATGACTATTGGGACCTTTGGGATTTGGGTGCC 840  
DB 781 CTGCTCACCCAGTGCACACACACCATCATGACTATTGGGACCTTTGGGATTTGGGTGCC 840  
QY 841 TACCTGCGAGGTGTGATACCATCTACTTAGCCAACTACACCTTCGGATTCCTCCGTTTC 900  
DB 841 TACCTGCGAGGTGTGATACCATCTACTTAGCCAACTACACCTTCGGATTCCTCCGTTTC 900  
QY 901 CTCAAAGTCTTTAAGCCAGAGGAGGAGGCTTCCTACCGGAATGGTGGGCGCATCCCTGCCAT 960  
DB 901 CTCAAAGTCTTTAAGCCAGAGGAGGAGGCTTCCTACCGGAATGGTGGGCGCATCCCTGCCAT 960  
QY 961 CTGTCCTCCACTCTTTAAGGCAITTAACACAGGCTGTCTCGGTCCACCTTCACCTCAAG 1020  
DB 961 CTGTCCTCCACTCTTTAAGGCAITTAACACAGGCTGTCTCGGTCCACCTTCACCTCAAG 1020  
QY 1021 GCAAAAGGAGTCACTTTGTTACGTGCGAGGAAGAGCCTTCGTATGGGAA 1068  
DB 1021 GCAAAAGGAGTCACTTTGTTACGTGCGAGGAAGAGCCTTCGTATGGGAA 1068  
RESULT 3  
AAC67965  
ID AAC67965 standard; cDNA; 1149 BP.  
XX AAC67965;  
AC AAC67965;  
XX  
DT 19-FEB-2001 (first entry)  
XX  
DE Rat hepatoma H35 cell alpha-2-fucosyltransferase cDNA.  
XX  
KW Rat; alpha-2-fucosyltransferase; cytostatic; neuroprotective;  
KW nootropic; gene therapy; Fucalalpha-2Galbeta1-3GalNAc; immunotherapy;  
KW immunosuppression; cancer; neurological disease;  
KW small cell lung carcinoma; ss.  
XX  
OS Rattus norvegicus.  
XX  
PN WO200064464-A1.  
XX  
PD 02-NOV-2000.  
XX  
XX 23-APR-1999; 99WO-US07384.  
XX  
XX 23-APR-1999; 99WO-US07384.  
XX  
XX (PACI-) PACIFIC NORTHWEST CANCER FOUND.  
XX  
XX Holmes EH, Sherwood AL;  
XX  
DR WPI; 2000-687262/67.  
DR P-PSDB; AAB36104.  
XX  
XX New rat ganglioside GM1-specific alpha-2-fucosyltransferase, useful for  
PT preparation of fucosyl GM1 which is useful as a nutritional composition  
PT or immunotherapeutic for cancer and neurological diseases -  
XX  
XX Claim 10; Fig 5; 91pp; English.  
XX  
XX The present sequence is given in a specification relating to a rat  
CC ganglioside GM1-specific alpha-2-fucosyltransferase protein. The protein  
CC or its cellular fraction is useful for synthesis of a molecule comprising  
CC Fucalalpha-2Galbeta1-3GalNAc, a glycolipid, glycoprotein, glycolipoprotein  
CC or a free oligosaccharide comprising Fucalalpha-2Galbeta1-3GalNAc.  
CC The method involves contacting alpha-2-fucosyltransferase with GDP-fucose  
CC and a molecule or glycolipid, glycoprotein, glycolipoprotein or  
CC oligosaccharide having a terminal Galbeta1-3GalNAc group. It is also  
CC useful for synthesis of fucosyl-GM1 by contacting the protein with

CC GDP-fucose and ganglioside GM1. The obtained glycoproteins,  
CC glycolipids and oligosaccharides are useful as  
CC nutritional compositions and fucosyl-GM1 is useful for inducing an  
CC immunotherapeutic or immunosuppressive action against cancer,  
CC neurological disease or small cell lung carcinoma.  
XX

SQ Sequence 1149 BP; 247 A; 341 C; 311 G; 250 T; 0 other;

Query Match 100.0%; Score 1068; DB 21; Length 1149;  
Best Local Similarity 100.0%; Pred. No. 8.9e-289;  
Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCACGACGGAATAGTGAAGCTCCAAACCCCTGTGACAGAGGAATATACCGATGACGACT 60  
DB 82 CTCACGACGGAATAGTGAAGCTCCAAACCCCTGTGACAGAGGAATATACCGATGACGACT 141  
QY 61 CAATGTCTCTGGGNAACACAGAAAGCCACAGATGCCAGCGACGACGACGATGGG 120  
DB 142 CAATGTCTCTGGGNAACACAGAAAGCCACAGATGCCAGCGACGACGACGATGGG 201  
QY 121 AATGGAGAGCTGCGGGGATGTTACGATCAATTCATTGGCGGCTGGGAACCCAGATG 180  
DB 202 AATGGAGAGCTGCGGGGATGTTACGATCAATTCATTGGCGGCTGGGAACCCAGATG 261  
QY 181 GCGGAATACGCCACACTCTTTGCACTGCGGATGAAACGACGCGCTTGCCTTCATCCCC 240  
DB 262 GCGGAATACGCCACACTCTTTGCACTGCGGATGAAACGACGCGCTTGCCTTCATCCCC 321  
QY 241 GCATCATGACACAGCTCTAGCCGCCATCTTCAGATCAGCCCTCCCGTGTACACAGC 300  
DB 322 GCATCATGACACAGCTCTAGCCGCCATCTTCAGATCAGCCCTCCCGTGTACACAGC 381  
QY 301 GACACGGCCAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGAGGT 360  
DB 382 GACACGGCCAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGAGGT 441  
QY 361 TACCGCCACATTCGGGACATTTGTGCGCTTCACGGATACCCGTGCTCCGTGACCTTC 420  
DB 442 TACCGCCACATTCGGGACATTTGTGCGCTTCACGGATACCCGTGCTCCGTGACCTTC 501  
QY 421 TACCACACCTGCGCCGACAGATCTGAGGAGTTCACCCGTGATGACACGACGCGGGAG 480  
DB 502 TACCACACCTGCGCCGACAGATCTGAGGAGTTCACCCGTGATGACACGACGCGGGAG 561  
QY 481 GAGGCCACAGCCCTTCTCGGTGCTGCGGATGAATGGGAGCCAGCAGTACTTTTGTG 540  
DB 562 GAGGCCACAGCCCTTCTCGGTGCTGCGGATGAATGGGAGCCAGCAGTACTTTTGTG 621  
QY 541 GGTGTCCATGTGCGCGAGGGGACTATGTGCATGTCTATGCTTAATGTGTGGAAGGCGTG 600  
DB 622 GGTGTCCATGTGCGCGAGGGGACTATGTGCATGTCTATGCTTAATGTGTGGAAGGCGTG 681  
QY 601 GTGGCTGACGGGGTACTCGAAAGCCCTGGATATGTTCCGGGACGCTATTCATCT 660  
DB 682 GTGGCTGACGGGGTACTCGAAAGCCCTGGATATGTTCCGGGACGCTATTCATCT 741  
QY 661 CCAGTCTTCTGTTTACAGCAACCGATGGCTGTGCGGAGAGCAATTAATGCTTCC 720  
DB 742 CCAGTCTTCTGTTTACAGCAACCGATGGCTGTGCGGAGAGCAATTAATGCTTCC 801  
QY 721 CGAGAGAGCTGGTGTTCGGGGCAATGATGTAGAGGGTGCACCAAGCAAGACTTCGGG 780  
DB 802 CGAGAGAGCTGGTGTTCGGGGCAATGATGTAGAGGGTGCACCAAGCAAGACTTCGGG 861  
QY 781 CTGCTCACCCACTGCAACACACCATCATGACTATTGGGACCTTGGGATTTGGCTGCC 840  
DB 862 CTGCTCACCCACTGCAACACACCATCATGACTATTGGGACCTTGGGATTTGGCTGCC 921  
QY 841 TACCTGGCAGGTGGTATACCATCTACTTAGCAACTACACCTTCCCGATTCCTCGGTC 900  
DB 922 TACCTGGCAGGTGGTATACCATCTACTTAGCAACTACACCTTCCCGATTCCTCGGTC 981  
QY 901 CTCGAAGTCTTTAAGCCAGAGCAGCGCTTCTACCCGGAATGGGTGGGCATCCCTCCCGAT 960

DB 982 CTCGAAGTCTTTAAGCCAGAGCAGCGCTTCTACCCGAATGGTGGGCATCCCTCCGAT 1041  
QY 961 CIGTCCCACTCCCTTAAGGCATTACACACGCGCTGTCTCGTCCGCTCCACTTCACCTCAAG 1020  
DB 1042 CTGTCCCACTCCCTTAAGGCATTACACACGCGCTGTCTCGTCCGCTCCACTTCACCTCAAG 1101  
QY 1021 GCAAAAGGAGTCACTTGTAGTTCGACGAGGAGCGCTTCTGATGGAA 1068  
DB 1102 GCAAAAGGAGTCACTTGTAGTTCGACGAGGAGCGCTTCTGATGGAA 1149

RESULT 4  
AAD27207

ID AAD27207 standard; DNA; 1149 BP.

XX AAD27207;

XX 09-APR-2002 (first entry)

XX Rat hepatoma H35 cell alpha1-2FucT DNA.

XX Rat: alpha1-2fucosyltransferase; alpha1-2FucT; antisense therapy;  
KW galactose beta1-3N-acetylglactosamine; Galbeta1-3GalNAc; glycolipid;  
KW glycoprotein; glycolipoprotein; oligosaccharide; fucosyl-GM1; cancer;  
KW gene therapy; oncogenic transformation; cytostatic; ganglioside; GM1;  
KW cell transformation; ds.

XX Rattus norvegicus.

XX Key Location/Qualifiers  
FH CDS 1..1143

FT /\*tag= a /product= "Rat hepatoma H35 cell alpha1-2FucT"

FT /\*tag= b /note= "Region which overlaps rat FTB"

FT US6329170-B1.

PN 11-DEC-2001.

PD 23-APR-1999; 990S-0298886.

XX 23-APR-1999; 990S-0298886.

XX (NWHO-) NORTHWEST HOSPITAL.

PI Holmes EH, Sherwood AL;

XX WPI; 2002-121132/16.

DR P-PSDB; AAE16622.

XX Rat hepatoma H35 cell alpha1-2fucosyltransferase, useful for producing  
PT GM1-specific alpha1-2fucosyltransferase enzyme by recombinant  
PT techniques and for detecting oncogenic transformation of test tissues -

XX Claim 1; Fig 5; 41pp; English.

XX The invention relates to rat GM1-specific alpha1-2fucosyltransferase  
CC (alpha1-2FucT) enzyme and its corresponding nucleic acid. This nucleic  
CC acid is specific for a carbohydrate moiety found in ganglioside GM1,  
CC a terminal galactose beta1-3N-acetylglactosamine (Galbeta1-3GalNAc)  
CC saccharide. Alpha1-2FucT DNA is useful for producing rat alpha1-2FucT  
CC protein by recombinant techniques. Alpha1-2FucT DNA is useful for the  
CC preparative synthesis of fucosyl containing glycolipids, glycoproteins,  
CC glycolipoproteins and oligosaccharide, and for preparing fucosyl-GM1.  
CC Alpha1-2FucT DNA is useful for detecting oncogenic transformation which  
CC involves assaying for changes in expression of alpha1-2 FucT. Since  
CC alpha1-2FucT is activated in cell transformation, antisense sequences  
CC derived from alpha1-2FucT DNA are useful for inhibiting, suppressing  
CC or treating cancer. Alpha1-2FucT DNA is useful in gene therapy and  
CC antisense therapy. The present sequence is rat hepatoma H35 cell

CC	alpha1-2Fuct DNA.	
XX	Sequence 1149 BP; 247 A; 341 C; 311 G; 250 T; 0 other;	
SQ	Query Match 100.0%; Score 1068; DB 24; Length 1149;	
	Best Local Similarity 100.0%; Pred. No. 8.9e-289;	
	Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 CTCAGCAGCAATAGTGAAGCTCCAAACCCCTGTGCAGAGAAGGAATTACCGATGACGACT 60	
DB	82 CTCACGACGCAATAGTGAAGCTCCAAACCCCTGTGCAGAGAAGGAATTACCGATGACGACT 141	
QY	61 CAAATGTCTCTCGGGAACACAGAAACCCAGAGATGCGACGGGACAGGAGCAGCATGGG 120	
DB	142 CAAATGTCTCTCGGGAACACAGAAACCCAGAGATGCGACGGGACAGGAGCAGCATGGG 201	
QY	121 AATGAGAGCTCCGGGGAATGTTACGATCAATTCATNTGGCCGCTGGGGAACAGATG 180	
DB	202 AATGAGAGCTCCGGGGAATGTTACGATCAATTCATNTGGCCGCTGGGGAACAGATG 261	
QY	181 GCGAATAGCCACACTCTTTCACTGGCCAGATGAACGGACGGCTTGCCTTCATCCCC 240	
DB	262 GCGAATAGCCACACTCTTTCACTGGCCAGATGAACGGACGGCTTGCCTTCATCCCC 321	
QY	241 GCATCCATGCACACAGCTCTAGCGCCCATCTTCAGGATCAGCCTCCCGGTGTTACACAGC 300	
DB	322 GCATCCATGCACACAGCTCTAGCGCCCATCTTCAGGATCAGCCTCCCGGTGTTACACAGC 381	
QY	301 GACAGGGCAAAAGATCCCATGGCAGATATACCATCTCAACGACTGGATGGAGAGGCT 360	
DB	382 GACAGGGCAAAAGATCCCATGGCAGATATACCATCTCAACGACTGGATGGAGAGGCT 441	
QY	361 TACCGCCACATTCGGGACACTTTGTGCTCTCAGGGATACCGGTCTCTCGACCTTC 420	
DB	442 TACCGCCACATTCGGGACACTTTGTGCTCTCAGGGATACCGGTCTCTCGACCTTC 501	
QY	421 TACCACACCTCGCCCGACAGATCTCTGAAGAGTTTCAACCTGCATGACACACGTGGGGAG 480	
DB	502 TACCACACCTCGCCCGACAGATCTCTGAAGAGTTTCAACCTGCATGACACACGTGGGGAG 561	
QY	481 GAGGCCAGGCTTCTGCTGCTGCTGGGGTGAATGGGAGCCAGCCGAGTACTTTTGTG 540	
DB	562 GAGGCCAGGCTTCTGCTGCTGCTGGGGTGAATGGGAGCCAGCCGAGTACTTTTGTG 621	
QY	541 GTGTCTCAGTTCGCGAGGGGACTATGTGCATGTATCCCTAATGTGGGAAGGCGGTG 600	
DB	622 GTGTCTCAGTTCGCGAGGGGACTATGTGCATGTATCCCTAATGTGGGAAGGCGGTG 681	
QY	601 GTGGCTGACCGGGGTTACCTGGAAGAGCCCTGGATATGTTCCGGCAGCGCTATTCACT 660	
DB	682 GTGGCTGACCGGGGTTACCTGGAAGAGCCCTGGATATGTTCCGGCAGCGCTATTCACT 741	
QY	661 CCAGTCTTCTGTTTACAGCAAGGTATGGCTGTGCGGGGAGAACATTAATGTTTCC 720	
DB	742 CCAGTCTTCTGTTTACAGCAAGGTATGGCTGTGCGGGGAGAACATTAATGTTTCC 801	
QY	721 CGAGGAGAGCTGTGTTCGCGGCAATGTTATGAGGGTTCGCGACCAAGGACTTCGGG 780	
DB	802 CGAGGAGAGCTGTGTTCGCGGCAATGTTATGAGGGTTCGCGACCAAGGACTTCGGG 861	
QY	781 CTGCTCACCAGTGCACCAACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCC 840	
DB	862 CTGCTCACCAGTGCACCAACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCC 921	
QY	841 TACCTGCGAGGTGGTATACCATCTACTTAGCAACATACACCTTCCGGATTTCCGGTTC 900	
DB	922 TACCTGCGAGGTGGTATACCATCTACTTAGCAACATACACCTTCCGGATTTCCGGTTC 981	
QY	901 CTCAAAGTCTTTAAGCCAGGAGCAGCTTCTACCCGAATGGGTGGCATCCCTGCGGAT 960	
DB	982 CTCAAAGTCTTTAAGCCAGGAGCAGCTTCTACCCGAATGGGTGGCATCCCTGCGGAT 1041	
QY	961 CTGTCCCACTCTTAAAGGCATTAACACCAGCCTGTCTCGGTCCCACTTCCACCTCAAG 1020	

DB	1042 CTGTCCCACTCCTTAAGGCATTAAACACAGCCTGTCTCCGTCCACTTCCACCTCAAG 1101	
QY	1021 GCAAAAGAGTGACACTTGTACGTCCAGGAAGAGCCTTCTGATGGAA 1068	
DB	1102 GCAAAAGAGTGACACTTGTACGTCCAGGAAGAGCCTTCTGATGGAA 1149	
RESULT 5		
AAV29003		
ID	AAV29003 standard; DNA; 1043 BP.	
XX	AC AAV29003;	
XX	DT 28-AUG-1998 (first entry)	
XX	DE Porcine secretor transferase (FUT2) gene.	
XX	KW Secretor; glycosyltransferase; FUT2; pig; epitope; antigen;	
KW	transgenic animal; xenotransplantation; organ transplant; ss.	
OS	Sus scrofa.	
XX	Key Location/Qualifiers	
FH	FT 9..1031	
FT	CDs /*tag= a	
XX	PN WO9807837-Al.	
XX	PD 26-FEB-1998.	
XX	PF 22-AUG-1997; 97WO-AU00540.	
XX	PR 23-AUG-1996; 96AU-0001823.	
XX	PA (AUST-) AUSTIN RES INST.	
XX	PI McKenzie IFC, Sandrin MS;	
XX	DR WPI; 1998-169148/15.	
XX	P-PSDB; AAW37855.	
PT	Nucleic acid encoding glycosyltransferase able to compete with	
PT	second such enzyme - particularly used to reduce expression of	
PT	unwanted carbohydrate epitope(s) on tissues intended for	
PT	transplantation	
XX	Claim 6; Fig 1A-B; 40pp; English.	
XX	This nucleotide sequence, designated clone pSe16.1, codes for	
CC	porcine secretor (Se) glycosyltransferase (see AAW37855), an enzyme	
CC	that has high affinity for type I and type III substrates. It was	
CC	isolated from a pig liver genomic library using full-length human	
CC	Se2 (FUT2) cDNA as probe. The nucleotide sequence shows about 36%	
CC	homology with human FUT1. Introduction of glycosyltransferase	
CC	nucleic acid, such as porcine Se nucleic acid, into a tissue	
CC	results in reduced expression of unwanted carbohydrate epitopes on	
CC	the tissue, especially porcine heart, liver, kidney or pancreas,	
CC	rendering it more suitable for transplantation, i.e. less	
CC	immunogenic and of increased immunological acceptability. A	
CC	claimed method of producing a cell from a donor species that is	
CC	immunologically acceptable to a recipient species involves reducing	
CC	levels of carbohydrate on the donor cell that causes it to be	
CC	recognised as non-self by the recipient by causing a nucleic acid	
CC	for a glycosyltransferase such as Se to be expressed in the cell.	
CC	Expression units, such as retroviral packaging or producer cells,	
CC	containing Se nucleic acids can be used in gene therapy.	
XX	SQ Sequence 1043 BP; 189 A; 352 C; 306 G; 196 T; 0 other;	
	Query Match 57.4%; Score 612.8; DB 19; Length 1043;	
	Best Local Similarity 80.6%; Pred. No. 2.3e-161;	
	Matches 716; Conservative 0; Mismatches 172; Indels 0; Gaps 0;	

```
QY 91 GAGATGCGACGGGACAGCGACAGCATGGGAATGGAGAGCTGCGGGGCATCTTCACGATC 150
DB 138 GTGACGGAGGTGACACAGAGAGCCCTCGAGCCCGAGCTGAAGGGCATGTGGACGATC 197
QY 151 AATTCCTATTGCGCGCTGGGGAACACAGATGGCGGAATACGCCACACTCTTTGCACCTGGCC 210
DB 198 AATGCCATCGCGCGCTGGGGAACACAGATGGGAGTACGCCACACCTGTACGCGCTGGCC 257
QY 211 AGGATGAACGACGGCTTGCCTTATCCCGCATCCATGCACACAGCTTCAGCGCCCATC 270
DB 258 AGGATGAACGCGCGCGCTTATCCCGCATCCATGCACACAGCTTCAGCGCCCATC 317
QY 271 TTCAGGATCAGCTTCCCGCTGTACACAGCGACAGCGCCAAAGATCCCATGGAGAAAT 330
DB 318 TTCAGGATCAGCTTCCCGCTGTACACAGCGACAGCGCCCGAGATCCCATGGAGAAAT 377
QY 331 TACCATCTCAACGACTGGATGGAGAGCGTTACCGCCACATTCGCGGACACTTTGTGCGC 390
DB 378 TACCACCTGAACGACTGGATGGAGAGCGTTACCGCCACATTCGCGGAGTACGTCGCGC 437
QY 391 TTCAGGATCAGCTTCCCGCTGTACACAGCGACAGCGCCAAAGATCCCATGGAGAAAT 450
DB 438 CTCAGGGCTACCGCTTCCCGCTGTACACAGCGACAGCGCCCGAGATCCCATGGAGAAAT 497
QY 451 GAGTTCAACCTGATGACACAGCGCGGAGGAGCGGCTTCTCGTGGTCTGCGG 510
DB 498 GAGTTCAACCTGATGACACAGCGCGGAGGAGCGGCTTCTCGTGGTCTGCGG 557
QY 511 GTGAATGGAGCGACCGAGTACTTTTGTGGGTGTCATGTGTCGCGGAGGGAGTATGTG 570
DB 558 GTGAATGGAGCGACCGAGTACTTTTGTGGGTGTCATGTGTCGCGGAGGGAGTATGTG 617
QY 571 CATGTCATCCCTAATGTGGAAGGCGTGTGGCTGACCGGGGTACTGTGAAGAGGCC 630
DB 618 CACGTGATGCCAAGTGTGGAAGGCGTGTGGCTGACCGGGGTACTGTGAAGAGGCC 677
QY 631 CTGGATATGTTCCGGGCGAGCTATCATCTCCAGTCTTGTGGTTTACAGCAACGATATG 690
DB 678 CTGGATGTTTCCGGGCGAGCTATCATCTCCAGTCTTGTGGTTTACAGCAACGATATG 737
QY 691 GCGTGGTCCGGGAGAAATTAATGCTTCCGAGAGAGCGTGTGTCGCGGGAATGTT 750
DB 738 GCGTGGTCCGGGAGAAATTAATGCTTCCGAGAGAGCGTGTGTCGCGGGAATGTT 797
QY 751 ATGAGGGGTCCGACGAGGAGTTCGCGTGTCTACCCAGTGCACACCAACCATCATG 810
DB 798 ATGAGGGGTCCGACGAGGAGTTCGCGTGTCTACCCAGTGCACACCAACCATCATG 857
QY 811 ACTATTGGGACCTTTGGGATTTGGCTGCTACCTGGCAGAGTGTGATACCATCTACTTA 870
DB 858 ACCATTGGCAGCTTCGGGATCTGGCGGCTACCTTGTGTTGGAGAGACCATCTACCTG 917
QY 871 GCGACTACACCTTCCGGATCTCGGTCTCCCAAGTCTTTAAGCCAGGAGCGGCTTC 930
DB 918 GCGAATACACGCTCCGGATCTCGGTCTCCCAAGTCTTTAAGCCAGGAGCGGCTTC 977
QY 931 CTACCGCAATGGTGGGACCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 978
DB 978 CTGCGCGAGTGGATGGGATCGAGGAGAGACCTGTCCCACTCTCTTAAG 1025
```

## RESULT 6

AAV21639

ID AAV21639 standard; cDNA; 1043 BP.

XX

AC

XX

XX

DT

08-JUL-1998 (first entry)

XX

DE

Pig secretor encoding cDNA.

XX

KW

Pig; secretor; chimeric; glycosyltransferase; gene therapy;

```
KW transplantation; ss.
XX
XX Sus scrofa.
XX
XX Location/Qualifiers
FH 9..1031
FT CDS
FT /*tag= a
FT /product= "secretor"
XX
XX WO9805768-A1.
XX
XX 12-FEB-1998.
XX
XX 01-AUG-1997; 97WO-AU00492.
XX
XX 21-AUG-1996; 96US-0024279.
XX 02-AUG-1996; 96AU-0001402.
XX
XX (AUST-) AUSTIN RES INST.
XX
XX McKenzie IPC, Sandrin MS;
XX
XX WPI: 1398-159170/14.
XX P-PSDB: AAW53101.
XX
XX Nucleic acids encoding chimeric glycosyltransferases - used for
altering carbohydrate levels on the surface of cells, useful in gene
therapy and transplantation
XX
XX Example 2; Fig 6; 51pp; English.
XX
XX The present sequence encodes pig secretor used in an example of the
present invention. The present invention describes nucleic acids (NA)
encoding a chimeric glycosyltransferase. The NAs comprise a catalytic
domain of a first glycosyltransferase (GT) and a localisation signal of
a second GT, whereby when the NA is expressed in a cell and where the
chimeric enzyme is located in an area of the cell where it is able to
complete for substrate with a second GT, resulting in reduced levels of
a product from the second GT. The NAs can be used to produce cells and
organs with desired glycosylation patterns. Products and methods of the
present invention can be used to reduce the levels of undesirable
epitopes in cells, tissues or organs which may be used in
transplantation or gene therapy..
XX
XX Sequence 1043 BP; 189 A; 352 C; 306 G; 136 T; 0 other;
```

```
Query Match 57.4%; Score 612.8; DB 19; Length 1043;
Best Local Similarity 80.6%; Pred. No. 2.3e-161;
Matches 716; Conservative 0; Mismatches 172; Indels 0; Gaps 0;
QY 91 GAGATGCGACGGGACAGCGAGCGATGGGAATGGAGAGCTGCGGGGCATGTTCACGATC 150
DB 138 GTGACGGAGGTGACACAGAGAGCCCTCGAGCCCGAGCTGAAGGGCATGTGGACGATC 197
QY 151 AATTCCTATTGCGCGCTGGGGAACACAGATGGCGGAATACGCCACACTCTTTGCACCTGGCC 210
DB 198 AATGCCATCGCGCGCTGGGGAACACAGATGGGAGTACGCCACACCTGTACGCGCTGGCC 257
QY 211 AGGATGAACGACGGCTTGCCTTATCCCGCATCCATGCACACAGCTTCAGCGCCCATC 270
DB 258 AGGATGAACGCGCGCGCTTATCCCGCATCCATGCACACAGCTTCAGCGCCCATC 317
QY 271 TTCAGGATCAGCTTCCCGCTGTACACAGCGACAGCGCCAAAGATCCCATGGAGAAAT 330
DB 318 TTCAGGATCAGCTTCCCGCTGTACACAGCGACAGCGCCCGAGATCCCATGGAGAAAT 377
QY 331 TACCATCTCAACGACTGGATGGAGAGCGTTACCGCCACATTCGCGGACACTTTGTGCGC 390
DB 378 TACCACCTGAACGACTGGATGGAGAGCGTTACCGCCACATTCGCGGAGTACGTCGCGC 437
QY 391 TTCAGGATCAGCTTCCCGCTGTACACAGCGACAGCGCCAAAGATCCCATGGAGAAAT 450
DB 438 CTCAGGGCTACCGCTTCCCGCTGTACACAGCGACAGCGCCCGAGATCCCATGGAGAAAT 497
```



QY	451	CAGTTACCCCTCATGACACAGTGGCGGAGGAGCCACAGGCTTCTCGGTGCTGGG	510
DB	498	GAGTTACCCCTCATGACACAGTGGCGGAGGAGCCACAGGCTTCTCGGTGCTGGG	557
QY	511	GTGAATGGAGACGACCGAGTACTTTGTGGTGTCCATGTGCGCGGAGGAGCTATG	570
DB	558	GTGAACGGAGCGGACCGAGTACTTTGTGGTGTCCATGTGCGCGGAGGAGCTATG	617
QY	571	CATGTATGCCCTAAATGTGTGGAAGGCGGTGGTGTACCGGGGTGTACCTGAAAGGCC	530
DB	618	CACGTGATGCCCAAGCTGTGGAAGGCGGTGGTGTACCGGGGTGTACCTGAAAGGCC	677
QY	631	CNGGATATCTCCGGGACGCTATTCATCTCCAGTCTTGGTGTACCAAGCAACGGTATG	690
DB	678	CNGGACTGTCTCCGGGACGCTATTCATCTCCAGTCTTGGTGTACCAAGCAACGGTATG	737
QY	691	GCCTGTGCGCGGAGAACATTAATGCTTCCCGAGGAGAGTGTGTCTCCGGCAATG	750
DB	738	GCCTGTGCTCGGAAACATCAATGCTTCCCGGCGGAGTGTGTCTCCGGCAATG	797
QY	751	ATGAGGGGTGCGCAGCAAGAGTCTCGGCTGCTCACCAGTGCRAACACACCATCATG	810
DB	798	ATCGAGGGTCTCCCGCAAGAGTCTCGGCTGCTCACCAGTGCRAACACACCATCATG	857
QY	811	ACTATTGGGACCTTTGGGATTTGGGCTGCTACCTGGCAGGTGGTGTATACCATCTACTTA	870
DB	858	ACATTTGGCAGTTCGGGATTCGGGCTGCTACCTGGCAGGTGGTGTATACCATCTACTTA	917
QY	871	GCAACTACACCTTCCTGGATTCCTGCTTCCCTCAAAAGTCTTTAAGCCAGAGGAGCTTC	930
DB	918	GCAAAATACAGCTCTCCGAGCTCTCCCTCAAAAGTCTTTAAGCCAGAGGAGCTTC	977
QY	931	CTACCCGAATGGGTGGGCTTCCCTGCGGATCTCTCCCTCAAAAGTCTTTAAG	978
DB	978	CTCCCGAGTGGATTTGGGATTCGGGCTGCTACCTGGCAGGTGGTGTATACCATCTACTTA	1025
RESULT 7			
AAV58323			
ID	AAV58323 standard; DNA; 2115 BP.		
XX			
AC	AAV58323;		
XX			
DT	20-NOV-1998 (first entry)		
XX			
DE	Human Sec2 coding sequence.		
XX			
KW	Sec2; alpha(1,2) fucosyltransferase; H blood group; secretor genotyping;		
KW	Gdp-L-fucose;beta-D-galactoside 2-alpha-L-fucosyltransferase; human;		
KW	FUT2; nonsecretor genotyping; ds.		
XX			
OS	Homo sapiens.		
XX			
XX			
FH	Key	Location/Qualifiers	
FT	CDS	64..1095	
FT		/*tag= a	
XX			
PN	US5807732-A.		
XX			
XX			
PD	15-SEP-1998.		
XX			
PF	28-FEB-1995; 95US-0395800.		
XX			
PR	28-FEB-1995; 95US-0395800.		
XX			
PA	(GIOR/) GIORGI D.		
PA	(KELL/) KELLY R J.		
PA	(LENN/) LENNON G.		
PA	(LOWE/) LOWE J B.		
PA	(ROUQ/) ROQUIER S.		
XX			
PI	Giorgi D, Kelly RJ, Lennon G, Lowe JB, Rouquier S;		

XX	WPI: 1998-520127/44.
DR	P-PSDB; AAM69332.
XX	
PT	DNA encoding fucosyltransferase enzyme - useful for producing
PT	recombinant enzyme and genotyping person as secretor or nonsecretor
XX	
PS	Claim 1: Column 45-50; 55pp; English.
XX	
CC	This sequence encodes the human Sec2 protein of the invention. The DNA
CC	encodes a alpha(1,2) fucosyltransferase and is the secretor
CC	alpha(1,2)fucosyltransferase locus, that cross hybridises with the
CC	H blood group alpha(1,2)fucosyltransferase gene. The DNA is useful for
CC	producing a recombinant human GDP-L-fucose:beta-D-galactoside
CC	2-alpha-L-fucosyltransferase (FUT2) which can be used for genotyping an
CC	individual as a secretor or nonsecretor as it is known that nonsecretors
CC	homozygous for a mutant allele of the FUT2 gene that has a stop codon in
CC	the position corresponding to amino acid 143.
XX	
SQ	Sequence 2115 BP; 505 A; 606 C; 552 G; 452 T; 0 other;
	Query Match 55.3%; Score 590.6; DB 19; Length 2115;
	Best Local Similarity 79.8%; Pred. No. 5e-155;
	Matches 709; Conservative 0; Mismatches 179; Indels 1; Gaps 1;
QY	128 AGCTGGGGGCGATGTTCCAGCATCAATTCATATGGCCGGCTGGGGAACACAGATGGCGCAAT 187
DB	239 AGCTCAGGGGATGTGGACGATCAATGCATAGCGCGCTGGGGAACACAGATGGCGGAGT 298
QY	188 ACGCCACACTCTTGCACCTGGCCAGGATGAACGAGCGCTTGGTTCATCCCGCATCCA 247
DB	299 ACGCCACACTGTACGCGCTGGCGAAGATGAACGGGCGCGCGCTTCATCCCGCCACAGA 358
QY	248 TGCACACAGCTCTAGCGCCCATCTTCAGGATCAGCCTCCGGGTCTTACACAGGACACAGG 307
DB	359 TGCACACAGCCTTGGCGCCCATCTTCAGAAATCACTTCGCGGTGCTGCACAGCGCACGG 418
QY	308 CCAAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATCGGAGGAGCGTTACCGCC 367
DB	419 CCAGCAGGATCCCTGGCAGAACTACCACTGACGACTGGATCGGAGGAGGAATACCGCC 478
QY	368 ACATTCGGGACACTTTGTGCGCTTCACGGGATACCGGTCTCTCTGACCTTCTTACCACC 427
DB	479 ACATTCGGGGGAGTACGTCCGCTTCACCGGTACCCCTGCTCTGACCTTCTTACCACC 538
QY	428 ACCTGGCCCGCAGAGATCCTGAAGSAGTTCACTCTGATGACACAGTGCAGGAGAGGCC 487
DB	539 ACCTCCCGCAGGAGATCCTCCAGAGTTCACCTTCGACGACCACTGCGGGAGAGGCC 598
QY	488 AGGCCTTCCTGCGTGTCTGCGGTGAATGGGAGCCAGCCAGTACTTTTGGGTGTC 547
DB	599 AGAAGTTCCTCGGGGCTGCAGTGAACGAGCCCGCGGACCTTTGTAGGGTCC 658
QY	548 ATGTGCGCGAGGAGACTATGTGATGTATGCTATGCTATGTTCCGGGACGCTATTCATCT 607
DB	659 ATGTGCGCGAGGAGACTATGTCCATGTATGCCAAAAGTGTGAAGGGGTGGTGGCG 718
QY	608 ACCGGGTTTACTCGGAAAGCCCTGGATATGTTCCGGGACGCTATTCATCTCCAGTCT 667
DB	719 ACCGGGATACCTACAGCGGCCCTGGACTGGTTCGAGCTCGCTACAGTCCCTCATCT 778
QY	668 TCGTGGTTACAAGCAACGGTATGGCCTGGTGGCGGAGAACATTAATGCTTCCGAGGAG 727
DB	779 TCGTGGTACACAGTAATGGCATGSCCTGGTGTGCGGAGAACATTAATGACACCTCCACGGTG 838
QY	728 ACGTGTGTTTCGGGGCAATGGTATGAGGGGTGCGCAGCCAGGACTTCGCGCTGCTCA 787
DB	839 ATGTGTGTTTGTCTGGGATGGCATTCAGGCTCACCTGCCAAAGATTTTGTCTTACTCA 898
QY	788 CCCAGTGCACACACCATCATGACTATTTGGGACCTTTGGGATTGGGCTGCTCCCTACCTGG 847
DB	899 CACAGTGAACACACCATCATGACCATTTGGAGATTCGGGATTCGGGCGGCATACCTCA 958



27-SEP-2000;	2000US-235840P.	XX
27-SEP-2000;	2000US-235863P.	XX
28-SEP-2000;	2000US-236028P.	XX
28-SEP-2000;	2000US-236032P.	XX
28-SEP-2000;	2000US-236033P.	XX
28-SEP-2000;	2000US-236034P.	XX
28-SEP-2000;	2000US-236109P.	XX
28-SEP-2000;	2000US-236111P.	XX
29-SEP-2000;	2000US-236842P.	XX
29-SEP-2000;	2000US-236891P.	XX
02-OCT-2000;	2000US-237172P.	XX
02-OCT-2000;	2000US-237278P.	XX
02-OCT-2000;	2000US-237294P.	XX
02-OCT-2000;	2000US-237316P.	XX
03-OCT-2000;	2000US-237425P.	XX
03-OCT-2000;	2000US-237598P.	XX
03-OCT-2000;	2000US-237604P.	XX
03-OCT-2000;	2000US-237606P.	XX
03-OCT-2000;	2000US-237608P.	XX
01-NOV-2000;	2000US-244867P.	XX
01-NOV-2000;	2000US-245084P.	XX
(AVAL-) AVALON PHARM.		XX
Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;		XX
Soppet DR, Weaver Z;		XX
WPI; 2002-188264/24.		XX
Screening for anti-neoplastic agent involves exposing cells to a		XX
chemical agent to be tested for anti-neoplastic activity, and		XX
determining a change in expression of a gene of a signature gene set		XX
Claim 1; SEQ ID 4648; 44pp; English.		XX
The present invention describes a method (M1) for screening for an		XX
anti-neoplastic agent. The method involves exposing cells to a chemical		XX
agent to be tested for anti-neoplastic activity, determining a change in		XX
expression of at least one gene (I) of a signature gene set, where (I)		XX
comprises a sequence (S) selected from 8447 sequences (given in ABL61664		XX
to ABL70110), or is at least 95% identical to (S), where a change in		XX
expression is indicative of anti-neoplastic activity. (I) has cytostatic		XX
activity and can be used in gene therapy. M1 can be used for screening		XX
an anti-neoplastic agent, and can be used for producing a product which		XX
is the data collected with respect to the anti-neoplastic agent as a		XX
result of M1, and the data is sufficient to convey the chemical		XX
structure and/or properties of the agent. M1 can be used in the		XX
treatment of cancer such as colon, breast, stomach, lung, thyroid,		XX
oesophageal, ovarian, kidney, prostate or pancreatic cancer.		XX
adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,		XX
infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine		XX
carcinoma, papillary carcinoma and Wilm's tumour.		XX
Sequence 3088 BP; 723 A; 839 C; 798 G; 728 T; 0 other;		XX
Query Match	55.2%; Score 589.4; DB 24; Length 3088;	XX
Best Local Similarity	78.7%; Pred. No. 1.2e-154;	XX
Matches	704; Conservative	XX
	0; Mismatches 191; Indels 0; Gaps 0	XX
QY	128 AGCTGGGGGCGATGTTCCAGATCAATTCATTCCTGCGGCTGGGAACACAGATGGCGAAT	187
Db	287 AGCTCAGGGGAGTGTGGACGATCAATGCAATAGGCGCGCTGGGAACACAGATGGCGGAGT	346
QY	188 AGCGCACACTCTTGCACTGGCCAGGATGAACGGACGGCTTGCGTTTCATCCCGCGATCCA	247
Db	347 ACGCCACACTGTACGCCCTGGCCAGATGACACGGCGCGCCGCTTCATCCCGGCCAGA	406
QY	248 TGCACACAGCTCTAGCGCCCATCTTCAGATCAGCCTCCCGGTGTTACACAGCAGCAGG	307
Db	407 TGCACAGCACCTGGCCCCCATCTTCAGAATCACCTCCCGTGTGTGACACAGCGCAGG	466



```
SQ Sequence 3088 BP; 723 A; 839 C; 798 G; 728 T; 0 other;
Query Match 55.2%; Score 589.4; DB 24; Length 3088;
Best Local Similarity 78.7%; Pred. No. 1.2e-154;
Matches 704; Conservative 0; Mismatches 191; Indels 0; Gaps 0;
QY 128 AGCTGGGGGATCTTCACGATCAATTCATGTCGCGCTGGGAGACAGATGGCGAAT 187
D 128 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 287 AGCTAGGGGATGTGGAGATCAATGATAGCGCGCTGGGAAACAGATGGCGAGT 346
QY 188 AGCCACACACTTTTGGCACTGGCCAGAGTAAAGCGCTTGGCTTCAATCCCGCATCCA 247
D 188 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 347 AGCCACACACTGTAGCGCTGGCCAAAGTAAAGCGCGCGCTTCAATCCCGCCAGA 406
QY 248 TGCACACAGCTCTAGCGCCCACTTGTAGGATCAACCTCCCGGTGTACACAGGCACAGG 307
D 248 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 407 TGCACACAGCCCTGGCGCCCATCTTCAGAAATCAACCTGCGCGGTGCTCACAGGCGCACGG 466
QY 308 CCAAAAAGATCCGATGSCAGAAATACCATCTCAAGGACTGATGGAGAGCGTTACCGCC 367
D 308 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 467 CCAGCAGATCCCTGGCAGACATCAACCTGACGACTGATGGAGAGGATACCGCC 526
QY 368 ACATTCGGGACACTTTGTGCGGTTCACGGGATACCGCTCTCTGTGACCTTCTACACC 427
D 368 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 527 ACTTCGGGGAGTACCTCCGCTTACCGCTACCGCTACCGCTCTCTGTGACCTTCTACACC 586
QY 428 ACCTGGCCCGAGATCCTGAAGAGTTTCACTGATGACACAGCTGCGGGAGAGGCC 487
D 428 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 587 ACCTCGCCCGAGAGATCTCCAGAGTTTCACTGATGACACAGCTGCGGGAGAGGCC 646
QY 488 AGCCCTCTCTCGCTGCTGCGGTGAATGGAGCGAGCTGATCTTTGTGGGTGTCC 547
D 488 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 547 AGAATTCCTGCGGGCTGACAGTGAACGGAGACGCGCGGACCTTTGTAGGGTCC 706
QY 548 ATGTGCGCGAGGGACTATGTGATGCTATGCTATATGTGTGAAGGGGTGCTGCTG 607
D 548 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 707 ATGTGCGCGAGGGACTATGTGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 766
QY 608 ACCGGGTATCCGGAAGAGCCCTGGATATGTCGGGACACCTATCTCTCAGTCT 667
D 608 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 767 ACCGGGATACCTACAGAGCCCTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 826
QY 668 TCCTGTTTACAAGCAACGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 727
D 668 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 827 TCCTGCTCACCAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 886
QY 728 ACCTGCTGCTGCGGCAATGTATGAGGGTGGCGGACCAAGGACTTGGCGCTGCTCA 787
D 728 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 887 ATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 946
QY 788 OCCAGTGCACACACACATCATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 847
D 788 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 947 CACAGTGTACACACACACATCATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1006
QY 848 CAGGTGCTGATACCACTACTTACCACTTACCACTTACCACTTACCACTTACCACTT 907
D 848 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1007 CGGGCGGAGACACATCTACCTGCGCCATACCACTTACCACTTACCACTTACCACT 1066
QY 908 TCTTTAAGCCAGAGGAGCGCTTCTTACCGAATGGGTGGGCACTCCGTGGGACTGTGCC 967
D 908 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1067 TCTTTAAGCCAGAGGAGCGCTTCTTACCGAATGGGTGGGCACTCCGTGGGACTGTGCC 1126
QY 968 CACTCTCTAAGGCAATTAACACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
D 968 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1127 CTTTACTCAAGCACTAATGCTGGCCCATCTTTTGAACCTTTTCTCTCTCTCTG 1181
```

RESULT 11

AAV21640

ID AAV21640 standard; cDNA; 1098 BP.

XX

AC AAV21640;

XX

```
DT 08-JUL-1998 (first entry)
XX Pig H transferase encoding cDNA.
DE DE DE DE DE DE DE DE DE DE DE DE DE DE DE DE DE DE DE DE DE DE DE
XX Pig H transferase; chimeric; glycosyltransferase; gene therapy;
KW KW KW KW KW KW KW KW KW KW KW KW KW KW KW KW KW KW KW KW KW KW KW
XX translocation; ss.
OS OS OS OS OS OS OS OS OS OS OS OS OS OS OS OS OS OS OS OS OS OS OS
XX Sus scrofa.
FH FH FH FH FH FH FH FH FH FH FH FH FH FH FH FH FH FH FH FH FH FH FH
FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT
FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT FT
XX /product= "H transferase"
PN PN PN PN PN PN PN PN PN PN PN PN PN PN PN PN PN PN PN PN PN PN PN
XX WO9805768-A1.
XX 12-FEB-1998.
PD PD PD PD PD PD PD PD PD PD PD PD PD PD PD PD PD PD PD PD PD PD PD
XX 01-AUG-1997; 97WO-AU00492.
PF PF PF PF PF PF PF PF PF PF PF PF PF PF PF PF PF PF PF PF PF PF PF
XX 21-AUG-1996; 96US-0024279.
PR PR PR PR PR PR PR PR PR PR PR PR PR PR PR PR PR PR PR PR PR PR PR
XX 02-AUG-1996; 96AU-0001402.
PA PA PA PA PA PA PA PA PA PA PA PA PA PA PA PA PA PA PA PA PA PA PA
XX (AUST-) AUSTIN RES INST.
PI PI PI PI PI PI PI PI PI PI PI PI PI PI PI PI PI PI PI PI PI PI PI
XX McKenzie IFC, Sandrin MS;
DR DR DR DR DR DR DR DR DR DR DR DR DR DR DR DR DR DR DR DR DR DR DR
XX WPI; 1998-159170/14.
XX P-PSDB; AAW53102.
XX Nucleic acids encoding chimeric glycosyltransferases - used for
PT altering carbohydrate levels on the surface of cells, useful in gene
PT therapy and transplantation
XX Example 3; Fig 7; 51pp; English.
CC The present sequence encodes pig H transferase used in an example of the
CC present invention. The present invention describes nucleic acids (NA)
CC encoding a chimeric glycosyltransferase. The NAs comprise a catalytic
CC domain of a first glycosyltransferase (GT) and a localisation signal of
CC a second GT, whereby when the NA is expressed in a cell and where the
CC chimeric enzyme is located in an area of the cell where it is able to
CC compete for substrate with a second GT, resulting in reduced levels of
CC a product from the second GT. The NAs can be used to produce cells and
CC organs with desired glycosylation patterns. Products and methods of the
CC present invention can be used to reduce the levels of undesirable
CC epitopes in cells, tissues or organs which may be used in
CC transplantation or gene therapy.
XX Sequence 1098 BP; 181 A; 367 C; 312 G; 238 T; 0 other;
```

```
Query Match 39.7%; Score 423.8; DB 19; Length 1098;
Best Local Similarity 59.6%; Pred. No. 1.9e-108;
Matches 591; Conservative 0; Mismatches 252; Indels 6; Gaps 1;
QY 136 GGCAATGTTACGATCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 195
D 136 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 235 GGGACCTGGACTATTTACCGGATGCGCGTGTGGGAACGAGATGGCGGATGCGCACG 294
QY 196 CTCTTTGCACTGGCAGGATGAACGAGCGCTTTCATTCATTCATTCATTCATTCATTC 255
D 196 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 295 CTCTGGCTTGGCGAGCTCAACGCGCGCGCGCTTATCCAGCTTCCATGCGCTGCGACGC 354
QY 256 GCTTACGCGCCCATCTTCAGGATCAGCGCTCCCGGTGTTCACAGCGCAGCGCGCAAAAG 315
D 256 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 355 GTCTGGCGCCCGGTGTTCCGATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414
QY 316 ATCCCATGGCAGAAATTACCATCTCAACGACTGATGAGAGCGGTTACCGCCCATTCGG 375
D 316 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 415 GCTCCTTGGCGGAGCTGGAGCTTCACGACTGGATGCTCGAGGATTATGCCCACTTAAG 474
QY 376 GGACACTTTGTGGCTTCACGGGATACCGGTGCTCTCTGACCTTCTACACGACCTTGGCG 435
D 376 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 475 GAGCCTGGCTGAAGCTCACGGCTTCCCTGCTCTGAGACTTCTCCACCAACCTCCGG 534
QY 436 CCAGAGATCCCTGAAGAGCTTACCTGATGACACAGTCCGGGAGGAGGCCAGCCCTTC 495
Db 535 GAGCAGATCCCGAGGAGTTCACCTGTCACGACACCTTCGGCAAGAGGCCAGGGGTA 594
QY 496 CTGGCT-----GGPCTCGGGTGAATGGAGGCCAGCCGAGTACTTTTGTGGGTCTCCAT 549
Db 595 CTGAGTCAGTCCGCTACCCCGCACAGGGACCCGCCAGCACACTTCGTGGGGTCCAC 654
QY 550 GTGGCCGAGGGGACTATGTGATGTCAATGCTATGTGTGAAGGGCGTGTGGCTGAC 609
Db 655 GTGCCCGCGGGAGTATCTCGTGTGATGCCAAGCGCTGSAAGGGGTGTGGGTGAC 714
QY 610 CGGGTTTACCTGGAAAGGCCCTGATATGTTCCGGGCACGCTATTCATCTCCAGTCTTC 669
Db 715 GCGGTTTACCTCCAGCAGGCTATGACATGTTCCGGGCCGATGACGAAGCCCGCTTT 774
QY 670 GTGGTTACAAGCAAGGATGACCTGTGCGGGGAGAACATTAATGCTTCCGAGGAGAC 729
Db 775 GTGGTCAACGACGACGGATGAGTGTGCGGGAAGACATCGACACCTCCCGGGGGAC 834
QY 730 GTGGTGTTCGGGGCAATGATATTAGGGGTGTCGCCAGCAAGGACTTCGCGTGTCTAC 789
Db 835 GTGATCTTTGTGCTGGCGATGGCGGGAGCCCGCCGAGGACTTTGCGCTGTGGTG 894
QY 790 CAGTGCACACCAACCAATCATGACTATTGGGACCTTTGGGATTTGGCTGCTCCACTGCGT 954
Db 895 CAGTGCACACCAACCAATCATGACTATTGGGACCTTTGGGATTTGGCTGCTCCACTGCGT 954
QY 850 GTGGTGTATACCATCTATTAGCAACATACACCTTCGCGGATTCCTGCTTCAAGTC 909
Db 955 GTGGGAGATACCATCTATTGGTGAATTCACCTTCGCCACCTTCAGCTTCCTGAATC 1014
QY 910 TTTAAGCAGAGGAGCGCTTCTACCGGATGGGTGGGATCCCTGCGATGTCGCCCA 969
Db 1015 TTTAAGCAGAGGCGCTTCTACCGGATGGGTGGGATCCCTGCGATGTCGCCCA 1074
QY 970 CTCCTTAAG 978
Db 1075 CTCAGATG 1083
```

## RESULT 12

AAI15872  
ID AAI15872 standard; DNA; 1269 BP.

AC AAI15872;

XX 13-MAY-1999 (first entry)

DE Fragment of pHT83xF encoding swine alpha(1,2) fucosyltransferase.

XX Swine alpha(1,2) fucosyltransferase; FUT1; Escherichia coli; resistant;  
KW E. coli-associated intestinal disorder; E coli infection; ss.

XX Synthetic.

OS Sus sp.

XX Key Location/Qualifiers

FT CDS 9..1106

FT /\*tag= a

XX W09853101-A2.

XX 26-NOV-1998.

XX 20-MAY-1998; 98WC-US10259.

XX 20-MAY-1997; 97US-0047181.

XX (BIOT-) BIOTECHNOLOGY RES & DEV CORP.

PI Bosworth BT;  
XX  
DR WPI: 1999-131692/11.  
DR P-PSDB; AAW97356.  
XX  
PT Identifying swine genetically resistant to E. coli associated  
PT diseases - using PCR-RFLP to assay for polymorphisms in the  
PT alpha(1,2) fucosyltransferase 1 gene  
XX  
PS Claim 6; Fig 1; 19pp; English.

XX The present sequence encodes swine alpha(1,2) fucosyltransferase  
CC (FUT1). The specification describes methods relating to Escherichia  
CC coli-resistant swine. One of the methods for identifying a swine  
CC resistant to E. coli-associated intestinal disorders, comprises  
CC determining whether the base at 307 of alpha(1,2) fucosyltransferase  
CC gene (FUT1) is adenine (sic), in which case the swine are resistant. The  
CC porcine FUT1 polymorphisms can be used to develop drugs for the  
CC treatment of swine having E. coli-associated disease. The methods can  
CC also be used in breeding programmes to identify swine with resistance  
CC to E coli infection.

XX Sequence 1269 BP; 219 A; 413 C; 362 G; 275 T; 0 other;

Query Match 39.7%; Score 423.8; DB 20; Length 1269;

Best Local Similarity 69.6%; Pred. No. 2e-108;

Matches 591; Conservative 0; Mismatches 252; Indels 6; Gaps 1;

```
QY 136 GCATGTTACAGATCAATTCATTTGGCGGTGGGAGACCATGAGGGAATACGCCACA 195
Db 243 GGGACTCGACTATTTACCGGATGGCGGTTTGGGAACCATGAGGATATCCACG 302
QY 196 CTTCTTGACTCGCCAGGATGAACGGGCTTTCATCCCTCCGATCCATGACACAC 255
Db 303 CTGCTGGCCCTGGCGAGCTCAACGGCCCGAGGCTTATCCAGCTCCATGACGCC 362
QY 256 GCTCTAGCGCCCATCTTTCAGGATCAGCTCCCGGTGTTACACAGGAGACAGCCAAAAG 315
Db 363 GTCCTGGCCCGCTGTTCCGCATCAGCTGCTCTCTGGCGCCCGAGGTAGAGGCAC 422
QY 316 ATCCCATGGCAGAAATACATCTCAACACTGGATGGAGGACGTTACGCCACATTCGG 375
Db 423 GCTCTTGGCGGAGCTGGAGCTTCACGACTCGATGTCGAGGATATGCCACTTAAG 482
QY 376 GGACACTTTGTCCGCTTTCACGGATACCGTCTGCTGACCTTCTACACACACCTGGC 435
Db 483 GAGCCCTGGCTGAAGCTACCGCTTCCCTGCTGAGACTTCTTCCACACCTCCGG 542
QY 436 CCAGAGATCCTGAAGGAGTTCACCTGTCATGACACAGTCCGGGAGGAGCCAGGCTTC 495
Db 543 GAGCAGATCCGACGAGTTCACCTTCACAGCACACCTTCGGCAAGAGGCCCGGGTA 502
QY 496 CTGGCT-----GGTCTGGGGTGAATGGAGCCAGCCAGTACTTTTGGGTGTCCAT 549
Db 603 CTGAGTCAGTTCGCTTACCCGACAGGGGACCGCCCGACACCTTCGTGGGGTCCAC 562
QY 550 GTGGCGCAGGGGACTATGTGATGCTCATGCTTAATGTGTGAAGGGCTGGTGGCTGAC 609
Db 663 GTGGCGCGGGGACTATGCTGATGCTCCAGCCGCTGAGAGGGGTGGTGGGTGAC 722
QY 610 CGGGGTACCTGGAAGGCCCTGATGATGTTCCGGGACGCTATTCATCTCAAGTCTTC 669
Db 723 GCGCGTTAATCCAGCAGCTATGGCTTCCGGGCGCGGATAGCAAGCCCGCTTT 782
QY 670 GTGGTTACAAGCAACGGTATGGCTTGGTCCGGGAGACATTAATGCTTCCCGAGGAC 729
Db 783 GTGGTCACCAAGCAACGGCATGGAGTGGTCCCGGAGAGACATCGACCTCCCGGGGGAC 842
QY 730 GTGGTGTTCGGGGCAATGATTTAGGGGTGTCGCCAGCAAGGACTTCGCGTGTCTACC 789
Db 843 GTGATCTTTGCTGGCGATGGGGGAGGCCCGCCCGCAGGACTTTGCGGTGCTGGTG 902
QY 790 CAGTGCACACCAACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCTACCTGSCA 849
```

Db 903 CASTGCAACACACCATCATGACCATGSCACCTTGGCTTCTGGSCCGCCACCTGGCT 962  
 QY 850 GTGTGTGATACCATCTACTTACCAACTACACCCCTTCCGGATTCTCCGTTCTCCTAAAGTC 909  
 Db 963 GTGTGAGATACCATCTACTTGTGCTAACTTCAACCTGCCCACTTCCAGCTTCTCTGAAGATC 1022  
 QY 910 TTAAAGCCAGAGCGAGCCCTTCTACCGAATAGGGTGGGCATCCCTGCCGATCTGTCCCA 969  
 Db 1023 TTAAACCGAGCGCTTCTCTGCCGAGTGGTGGGCATTAATGACAGACTTGTCTCCA 1082  
 QY 970 CTCCTTAAG 978  
 Db 1083 CTCAGATG 1091

## RESULT 13

AAX03811  
 ID AAX03811 standard: DNA; 1269 BP.

XX AAX03811;

DT 01-APR-1999 (first entry)

XX Swine alpha-1,2-fucosyltransferase 1 encoding DNA.

KW Swine; pig; alpha-1,2-fucosyltransferase 1; FURL1; resistance;  
 KW Escherichia coli; infection; oedema; postweaning diarrhoea;  
 KW intestinal disorder; polymorphism; ss.

XX Sus scrofa.

FH Key Location/Qualifiers  
 FT CDS 9..1106  
 FT /\*tag= a

XX W09853102-A1.

XX 26-NOV-1998.

XX 20-MAY-1998; 98WO-US10318.

XX 20-MAY-1997; 97US-0047181.

PA (BIOT-) BIOTECHNOLOGY RES & DEV CORP.  
 PA (SWIT-) SWISS FEDERAL INST TECHNOLOGY ZURICH.  
 PA (USDA ) US SEC OF AGRIC.

XX Bosworth BT, Vogeli P;

XX WPI; 1999-059746/05.

XX P-PSDB; AAW30630.

XX New method of identifying swine that are resistant to intestinal  
 PT colonisation by Escherichia coli - comprises use of genetic  
 PT polymorphic markers, used for breeding swine resistant to  
 PT Escherichia coli-related diseases

XX Claim 6; Fig 1; 35pp; English.

CC A method has been developed for the identification of swine that are  
 CC resistant to intestinal colonisation by E. coli. The method comprises  
 CC determining whether a genetic polymorphism associated with resistance to  
 CC colonisation is present in a swine sample, and then inferring that the  
 CC swine is resistant if it is homozygous for the polymorphism. The method  
 CC uses the swine alpha-1,2-fucosyltransferase (FURL1) polymorphism. The  
 CC present sequence encodes swine FURL1. The method enables the breeding of  
 CC swine that are resistant to E. coli-related diseases. This method  
 CC comprises breeding swine that have a genetic polymorphism in the FURL1  
 CC gene. More particularly, the identification method identifies swine that  
 CC are resistant to E. coli-related intestinal disorders if, in a sample  
 CC taken, the only nitrogen base at residue 307 in the FURL1 gene is adenine.  
 CC Larger amplified fragments from the assay can be used for RFLP analysis,

CC and the assay itself is used as a basis for a kit, applied to swine of  
 CC any age, in detecting polymorphisms associated with E. coli F18  
 CC receptors. The polymorphisms are useful in developing drugs to treat  
 CC swine with E. coli-related diseases. However, a mutated form of the  
 CC porcine FURL1 gene may interfere with the normal enzyme and prevent it  
 CC from producing the intestinal receptor for F18. The detection of  
 CC polymorphic markers in the method disclosed enables the detection and  
 CC treatment of E. coli-related intestinal diseases in swine, where there  
 CC has been no success using antibiotics due to unsuccessful prophylaxis.

XX Sequence 1269 BP; 219 A; 413 C; 362 G; 275 T; 0 other;

Query Match 39.7%; Score 423.8; DB 20; Length 1262;

Best Local Similarity 69.6%; Pred. No. 2e-108;

Matches 591; Conservative 0; Mismatches 252; Indels 6; Gaps 1;

QY 136 GGCATGTTCCAGCATCAATTCATTTGGCGGCTGGGAACAGATGGCGGAATACGCCACA 195  
 Db 243 GGGACCTGGACTATTTACCCGGATGSCCGTTTGGCAACNATGGACAGATGCCACG 302  
 QY 196 CTCCTTGGCACTGGCCAGGATGAACGAGCGCTTGCCTTATCCCCCGCATCCATPGCAAC 355  
 Db 303 CTCGTGGCCCTGGCGAGCTCAACGGCGCCAGGCTTCATCCAGCTGCCATGCCACGCC 362  
 QY 256 GCTCTAGCGCCCATCTTCAGGATCAAGCTCCCGGTGTACACAGCGACACAGGCCCAAAAG 315  
 Db 363 GTCCCTGGCCCGCTGTTCGGCATCACCTGCTCCGTCTGGCGCCCGAGGATACAGCGCAC 422  
 QY 316 ATCCCATGGCAGAAATTAACCATCTCAACGACTGATGGAGAGCGTTTACCGGCCACATTCGG 375  
 Db 423 GCTCTTGGCGGAGCTGGAGCTTCACGACTGATGTCGGAGGATTATGCCCATTAAG 482  
 QY 376 GGACACTTTGTGGCTTTCACGGGATACCGTCTCTGGACCTTCTACACACACCTGGCGC 435  
 Db 483 GAGCCCTGGCTGAAGCTCACCGCTTCCCGCTTCCCTGCTCTGGACCTTCTCCACCACTCCGG 542  
 QY 436 CCAGAGATCTTGAAGGAGTTCAACCTGTCATGACACGTGCGGAGGAGGCCAGGGCTTC 495  
 Db 543 GAGCAGATCCGAGCGAGTTTCAACCTTGCACGACACCTTTCGGCAAGAGGCCCAAGGGGTA 602  
 QY 496 CTGCGT-----GGTCTGGCGGTGAATGGGAGCAGCCGAGTACTTTTGGCTGTCCAT 549  
 Db 603 CTGAGTCAGTTCCGTCTACCCGCGCACAGGGGACCGCCGCCAGCACCTTCGTGGGGTCCAC 662  
 QY 550 GTGCGCCGAGGGGACTATGTGATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 609  
 Db 663 GTGCGCCGCGGGGACTATCTGCTGTGATGCCCAAGCGCTGGAAGGGGGTGGTGGTGAC 722  
 QY 610 CGGGGTACCTGGAAAAGGCCCTGGATATGTTCCGGGCGACGCTATTTCATCTCCAGTCTTC 669  
 Db 723 GGCCTGTACCTCCAGCAGGCTATGGACTGTTCCGGGGCCCGATACGAAGCCCGCTCTTT 782  
 QY 670 GTGGTTACAAGCAACGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 729  
 Db 783 GTGGTCAACCAACCGCATGAGTGTGCTCCGGAAGAACATCGACACCTCCCGGGGGGAC 842  
 QY 730 GTGGTGTCCGGGCAATGTTATGAGGGGTGCCAGGCCAAGGACTTCGGCTGCTCAACC 789  
 Db 843 GTGATCTTTGCTGGCGATGGGGGGGCGCGCGCCCGAGGAGCTTTGGCTGCTGGTG 902  
 QY 790 CAGTCAACACACACCATCATGACTATTGGGACTTTGGGATTTGGCTGCTGCTGCTGCTGCTGCTG 849  
 Db 903 CAGTCAACACACCATCATGACTATTGGGACTTTGGGACTTTGGGCTTGGGGCCCTTACCTGGCT 962  
 QY 850 GTGGTGTATACCATCTACTTACCAACTACACCTTCCGGATTCTCCGTTCTCCTCAAGATC 909  
 Db 963 GTGGGAGATACCATCTACTTGGTAACTTCAACCTCCCGACCTTCCAGCTTCTCGAATC 1022  
 QY 910 TTAAAGCCAGAGCGACCTTCTTACCCGAATGGGTGGGCAATCCCTGCCGATCTGTCCCA 969  
 Db 1023 TTAAACCGGAGCTGCCCTTCTGCGCGAGTGGTGGCATTAATGACAGACTTGTCTCCA 1082  
 QY 970 CTCCTTAAG 978

```

Db      1083 CTCACGATG 1091
|||||
RESULT 14
AAZ94417
ID AAZ94417 standard; DNA; 1269 BP.
XX
AC AAZ94417;
XX
XX
DT 18-JUL-2000 (first entry)
XX
DE Pig alpha-1-2 fucosyltransferase FUT1 gene.
DE
KW Alpha-1-2 fucosyltransferase; FUT1; pig; polymorphism;
KW Escherichia coli; resistance; ss.
XX
OS Sus scrofa.
XX
FH Key Location/Qualifiers
FT CDS 9..1106
FT /tag= a
FT /transl_except= (pos:726..728, aa:Ala)
FT variation replace(315,A)
FT /tag= b
FT /note= "adenine is substituted for guanine in
FT resistant pigs"
XX
XX WO200016641-A1.
XX
XX 30-MAR-2000.
XX
XX 17-SEP-1999; 99WO-US21408.
XX
XX 18-SEP-1998; 98US-0151592.
XX
XX (BIOT-) BIOTECHNOLOGY RES & DEV CORP.
XX (USDA ) US SEC OF AGRIC.
XX
XX Bosworth BT, Ridpath J, Wiseman B;
XX
XX WPI; 2000-283404/24.
XX
XX P-PSDB; AAY79302.
XX
XX Improving weight gain in swine using swine genetically resistant
XX Escherichia coli and feeding swine high levels of plant based protein
XX
XX Disclosure; Fig 1; 33pp; English.
XX
XX This is the nucleotide sequence of the pig FUT1 gene encoding
XX alpha-1-2 fucosyltransferase (see AAY79302). A polymorphism at
XX position 307 of the coding region is associated with susceptibility
XX to F18 Escherichia coli colonization; pigs homozygous for adenine
XX at position 307 are resistant to colonization, while heterozygous
XX animals and animals homozygous for guanine at position 307 are
XX susceptible to colonization. A claimed method for improving weight
XX gain in pigs involves selecting animals that are genetically
XX resistant to E. coli colonization and feeding these animals high
XX levels of plant-based protein concentrate. A claimed method for
XX preventing F18 E. coli colonization in swine, especially swine that
XX are genetically susceptible to F18 E. coli colonization, involves
XX replacing some or all of the plant-based proteins in the diet with
XX animal-based proteins. The polymorphism in the FUT1 gene is also
XX useful for developing drugs to treat swine that have E. coli
XX associated disease. The polymorphism can be detected using
XX PCR-RFLP tests (see also AAZ94418-19).
XX
XX Sequence 1269 BP; 219 A; 413 C; 362 G; 275 T; 0 other;
XX
XX Query Match 39.7%; Score 423.8; DB 21; Length 1269;
XX Best Local Similarity 69.6%; Pred. No. 2e-108;
XX Matches 591; Conservative 0; Mismatches 252; Indels 6; Gaps 1;

```

```

QY 136 GGCATGTTCAAGATCAATTCATTTGCCCGCTGGGAAACAGATGGCGAAATACGCCACA 195
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 243 GGGACCTGGACTATTTACCCGGATGGCCGGTTGGGAACACAGATGGAGATGCCAGC 302
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 196 CTCCTTGCACTGCCAGGATGAACGACGGCTTGGCTTCATCCCCCATCCATGCACAAC 255
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 303 CTGCTGGCCCTGGCGCAGCTCAACGGCCGCCAGGCTTCATCCAGCTCCCATGCAAGCC 362
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 256 GCTCTAGCGCCCATCTTCAGGATCAGCTCCCGGTGTTACAGCAGACACGGCCCAAAAG 315
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 363 GTCCTGGCCCGCGTGTTCGGCATCAGCTGCTGCTGCGCGCCGAGGTACAGGAC 422
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 316 ATCCCATGGCAGAATTACCATCTCAAGCATGATGGAGAGCGTTACCGCCACATTCG 375
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 423 GCTCCTTGGCGGAGCTGGAGCTCAGACTGATGTCGGAGATTATGCCCACTTAAG 482
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 376 GGACACTTTGTGCGCTTCACGGGATACCGGTCTCTCTGGACCTTCTACCAACCTGGCC 435
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 483 GAGCCCTGGCTGAAGCTCACCGCTTCCCTCTCTCTGGACCTTCTTCCACCACTCCGG 542
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 436 CCAGAGATCCTGAAGAGTTTACCCCTGATGACCATGCGGGAGAGCCCGACCTTTC 495
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 543 GAGCAGATCCGACGCGAGTTTACCCCTGACAGCACCTTCGGCAAGAGCCCGAGGGGTA 602
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 496 CTGCGT-----GGTCTGCGGTGAATGGGAGCCAGCGAGTACTTTTGTGGGTCTGCAT 549
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 603 CTGAGCTCAGTTCCGTCTACCCCGCACAGGGNACCCCGACACCTTCGTGGGGTCCAC 662
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 550 GTGCCCGAGGGGACTATGTGATGTCATGCTAAATGTGTGSAAGGGGTGGTGGCTGAC 609
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 663 GTGCCCGCGGGGACTATCTGCTGTGATGCCAAGCGCTGSAAGGGGTGGTGGGTGAC 722
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 610 CGGGTTTACCTGGAAAGGCCCTGATGTAIGTCCGGGACAGCTATTCATTCACCTTTC 669
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 723 GCGCGTTTACCTCCAGCAGGCTATGAGCTGGTTCCGGGCCGCGATAGCAAGCCCGCTCTT 782
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 670 GTGTTACAAGCAACGGTATGSCCTGGTCCGGGAGAACATTAAATGCTTCCGAGGAGAC 729
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 783 GTGTTACCAAGCAACGGTATGAGTGGTCCGGGAGAACATCGACACTCCCGGGGGAC 842
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 730 GTGGGTTCGGGGCAATGGTATGTAGGGGTGGCCAGCCAAAGACTTCGCGTGTCTACC 789
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 843 GTGATCTTTGTGCGGATGGCGGAGCGCGCGCCCGCAGGGACTTTCGCTGTGTGTG 902
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 790 CAGTCAACCAACCATCATCATCTATTCGGACCTTGGGATTTGGGCTGCGCTACCTGCA 849
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 903 CAGTCAACCAACCATCATCATCTATTCGGACCTTGGGCTTCTGGGGCGCGCTACCTGCT 962
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 850 GGTGTGATACCATCTACTTAGCCAACTACACCTTCGGGATTCCTCCCTCAAGTTC 909
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 953 GGTGGAGATACCATCTACTTAGCTAACTTACCTTCACCTGCCACTTCACGCTCTCTGAAGATC 1022
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 910 TTTAAGCCAGAGGAGCGCTTCTACCCCAATGGGTGGGCAATCCCTGCGCATCTGTCCCA 969
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1023 TTTAAGCCAGAGGAGCGCTTCTACCCCAATGGGTGGGCAATTAATGACAGACTTGTCTCCA 1082
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 970 CTCCTTAAG 978
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1083 CTCAGATG 1091
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

## RESULT 15

AAI72831

ID AAI72831 standard; cDNA; 1269 BP.

XX

AC AAI72831;

XX

DT 22-JUL-2002 (first entry)

XX

DE FUT1 cDNA.

XX

KW Gene: pig; swine; alpha (1,2) fucosyltransferase; FUT1; weight gain;



Fl8; Escherichia coli; ss.  
Sus scrofa.  
Key Location/Qualifiers  
CDS 9..1106  
/\*tag= a  
/product= "FUT1"  
315  
mutation  
/\*tag= b  
/phenotype= "Confers resistance to Fl8 E. coli"  
US6355859-B1.  
XX 12-MAR-2002.  
XX 18-SEP-1998; 98US-0151592.  
XX 20-MAY-1997; 97US-047181P.  
XX (BIOT-) BIOTECHNOLOGY RES & DEV CORP.  
XX (USDA ) US SEC OF AGRIC.  
XX Bosworth B, Ridpath J, Wiseman B;  
XX WPI: 2002-391652/42.  
XX P-PSDB; AAB47995.  
XX Feeding swine genetically susceptible to Fl8 Escherichia coli  
XX colonization with a low plant protein based diet increases weight gain  
XX and lowers intestinal disease associated with E. coli infection  
XX Claim 1; Column 13-18; 9pp; English.  
XX This sequence represents the swine alpha (1,2) fucosyltransferase  
XX (FUT1) gene. A FUT1 gene in which there is a base other than adenine  
XX at position 307, may be used for improving weight gain in swine that  
XX are genetically susceptible to Fl8 Escherichia coli. The weight gain  
XX may be activated by feeding a diet of at least 40% animal based  
XX proteins. The feeding method is used to control Fl8 E. coli associated  
XX intestinal disease in swine.  
XX  
XX Sequence 1269 BP; 219 A; 414 C; 361 G; 275 T; 0 other;  
Query Match 39.7%; Score 423.8; DB 24; Length 1269;  
Best Local Similarity 69.6%; Pred. No. 2e-108;  
Matches 591; Conservative 0; Mismatches 252; Indels 6; Gaps 1;

136 GGCATGTTACAGTCAATTCACATGCGCGCTGGGAACAGATGGGCGAATACGCCACA 195  
196 CTCCTTGGACIGGCAGGATGAACGACGCTTGCCTTCATCCCGCATCCATCCACAC 255  
303 CTGCTGGCCCTGGCGAGTCAACGCCGCCAGGCGCTTATCCAGCTGCCATGCGACGCC 362  
256 GGTCTAGCGCCCATCTTACAGTACGCTCCCGGTGTTACACAGCGACACGCGCAAAAG 315  
363 GTCTGGCCGCCGTGTTCCGATCAAGCTGCGCTTCCTGGCGCCCGAGGTAGACAGGCAC 422  
316 ATCCCATGGCAGAAATACCATCTACAGCTGGATGGAGAGCGGTTACCGCCACATCCG 375  
423 GTCCTTGGCGGAGTGGAGCTTCAGACTGGATGCGGAGGATATGCCACTTAAAG 482  
376 GGACACTTTTGGCGCTACGGGATACCGCTGCTCGTGGACCTTCTACCAACACCTGGCG 435  
483 GAGCCCTGGCAGCTACCGGCTTCCCTGCTGCTGAGACCTTCTCCACACCTCCCG 542  
436 CCAGAGATCCTGAGGAGTCAACCTGCGATGACCACTGCGGAGGAGGCCACAGGCTTC 495  
543 GAGCAGATCCGCGAGGAGTCAACCTGCGAGCACCACCTTCGCGAAGGCCCCAGGGGGA 602  
496 CTGCGT- - - - -GGTCTGCGGGTGAATGGAGGCCAGCCAGTACTTTTGTGGGTGTCCAT 549

Search completed: May 26, 2003, 12:08:26  
Job time : 278.996 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 05:50:26 ; Search time 58.2896 Seconds  
(without alignments)

5619.028 Million cell updates/sec

Title: US-10-040-863-9

Perfect score: 1068

Sequence: 1 cttccagcagcgaatagttaa.....gaagagccttctgtatgggaa 1068

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1068	100.0	1068	4	US-09-298-886-9
2	1068	100.0	1149	4	US-09-298-886-7
3	612.8	57.4	1043	4	US-09-254-077A-5
4	590.6	55.3	2115	1	US-08-395-800A-7
5	516.6	48.4	1144	1	US-08-395-800A-1
6	423.8	33.7	1269	4	US-09-151-592-1
7	408.2	38.2	1155	1	US-08-208-889A-1
8	408.2	38.2	1155	2	US-08-433-271-1
9	405	37.9	999	5	PCT-US91-00899-10
10	405	37.9	1136	1	US-08-395-800A-9
11	405	37.9	1174	5	PCT-US95-07554-3
12	405	37.9	1199	1	US-08-395-800A-5
13	405	37.9	2268	4	US-08-675-773B-4
14	405	37.9	3373	1	US-08-273-411-2
15	405	37.9	3791	4	US-08-675-773B-3
16	405	37.9	8174	1	US-07-914-281-5
17	405	37.9	8174	1	US-08-393-246-5
18	405	37.9	8174	1	US-08-525-058A-5
19	405	37.9	8174	2	US-08-696-731-5
20	405	37.9	8174	4	US-08-042-531-5
21	405	37.9	8174	5	PCT-US91-00899-3
22	403.4	37.8	1155	1	US-08-434-151-1
23	403.4	37.8	1155	2	US-08-715-259-1
24	72.8	6.8	100	4	US-09-298-886-29
25	71.2	6.7	100	4	US-09-298-886-23
26	71.2	6.7	100	4	US-09-298-886-26
27	69.6	6.5	100	4	US-09-298-886-20

#### ALIGNMENTS

##### RESULT 1

US-09-298-886-9  
; Sequence 9, Application US/09298886  
; Patent No. 6329170  
; GENERAL INFORMATION:  
; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
; FILE REFERENCE: 8511-029  
; CURRENT APPLICATION NUMBER: US/09/298,886  
; CURRENT FILING DATE: 1999-04-26  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 1068  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1062)  
US-09-298-886-9

Query Match 100.0%; Score 1068; DB 4; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 2.2e-305;  
Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CTCCAGCAGCGAATAGTGAAGCTCCAAACCCCTGTGAGAGGAAGGAATACCGATGAGGACT	60
DB	1	CTCCAGCAGCGAATAGTGAAGCTCCAAACCCCTGTGAGAGGAAGGAATACCGATGAGGACT	60
QY	61	CAATGTCTCGGGAACACAGAAAGCCAGAGATGCGAGGGAGCAGGAGCAGCATGGG	120
DB	61	CAATGTCTCGGGAACACAGAAAGCCAGAGATGCGAGGGAGCAGGAGCAGCATGGG	120
QY	121	ATGAGAGAGCTCGGGGGCATGTTCAGGATCAATTCATTTGCCCGCTGGGGAACAGATG	180
DB	121	ATGAGAGAGCTCGGGGGCATGTTCAGGATCAATTCATTTGCCCGCTGGGGAACAGATG	180
QY	181	GGCGAATAGCCACACTCTTTGCACTGGCCAGGATGAACGGACGCTTGCCTTCATCCCC	240
DB	181	GGCGAATAGCCACACTCTTTGCACTGGCCAGGATGAACGGACGCTTGCCTTCATCCCC	240
QY	241	GCATCCATGCACAAAGCTTCTACGCCCATCTTCAGGATCAGCCCTCCCGGTGTTACACAGC	300
DB	241	GCATCCATGCACAAAGCTTCTACGCCCATCTTCAGGATCAGCCCTCCCGGTGTTACACAGC	300
QY	301	GACACGGCCAAAGATCCCATGGAGATGATGAGGAGGAGGCT	360
DB	301	GACACGGCCAAAGATCCCATGGAGATGATGAGGAGGAGGCT	360

361	QY	TACCGCCACATTCCGGGACATTTGTGCGCTTACGGGATACCCGTGCTCGTGGACCTTC	420
361	Db	TACCGCCACATTCCGGGACATTTGTGCGCTTACGGGATACCCGTGCTCGTGGACCTTC	420
421	QY	TACCACCACCTGGCCACAGATCCTGANGAGTTCACCTGCATCACCACGTGCGGGAG	480
421	Db	TACCACCACCTGGCCACAGATCCTGANGAGTTCACCTGCATCACCACGTGCGGGAG	480
481	QY	GAGGCCCCAGCGCTTCCTGCGTGGTCTGCGGGTGAATGGGAGCAGCGGACTTTTGTG	540
481	Db	GAGGCCCCAGCGCTTCCTGCGTGGTCTGCGGGTGAATGGGAGCAGCGGACTTTTGTG	540
541	QY	GTGTGTCATGTGCGCGGAGGGGACTATGTGCATGTATGCCCTAATGTGTGGAAGGCGTG	600
541	Db	GTGTGTCATGTGCGCGGAGGGGACTATGTGCATGTATGCCCTAATGTGTGGAAGGCGTG	600
601	QY	GTGCGCTGACCGGGGTACCTGGAAAGGCCCTGGATATGTCGGGACAGCCTATTCATCT	660
601	Db	GTGCGCTGACCGGGGTACCTGGAAAGGCCCTGGATATGTCGGGACAGCCTATTCATCT	660
661	QY	CCAGTCTTCGTGGTTACAAGCAACGGTATGGGCTGGTGCGGGAGAACATTAATGCTTCC	720
661	Db	CCAGTCTTCGTGGTTACAAGCAACGGTATGGGCTGGTGCGGGAGAACATTAATGCTTCC	720
721	QY	CGAGGAGACGTGGTGTTCGCGGGCAATGTATTTAGGGGTGCGCAGCAAGACTTCGGG	780
721	Db	CGAGGAGACGTGGTGTTCGCGGGCAATGTATTTAGGGGTGCGCAGCAAGACTTCGGG	780
781	QY	CTGCTCACCCAGTGCAAACACACACCATCANGACTATTGGGACCTTGGGATTTGGGCTGCC	840
781	Db	CTGCTCACCCAGTGCAAACACACACCATCANGACTATTGGGACCTTGGGATTTGGGCTGCC	840
841	QY	TACCTGGCAGSGTGGTGATACCATCTACTTTAGCCAACTACACCCCTCCGGATTCTCGTTC	900
841	Db	TACCTGGCAGSGTGGTGATACCATCTACTTTAGCCAACTACACCCCTCCGGATTCTCGTTC	900
901	QY	CTCAAAGTCTTTAAGCCAGAGCAGCCTTCCTACCCGAATGGTGGGCATCCCTGCGGAT	960
901	Db	CTCAAAGTCTTTAAGCCAGAGCAGCCTTCCTACCCGAATGGTGGGCATCCCTGCGGAT	960
961	QY	CTGTCCCCACTCCTTAAGGCATTAAACACAGCCTGTCTCGGTGCCACTTCCACCTCAAG	1020
961	Db	CTGTCCCCACTCCTTAAGGCATTAAACACAGCCTGTCTCGGTGCCACTTCCACCTCAAG	1020
1021	QY	GCAAAAGAGTCACTTGTGTACGTGCGAGGAAGCCTTTGTATGGGAA	1068
1021	Db	GCAAAAGAGTCACTTGTGTACGTGCGAGGAAGCCTTTGTATGGGAA	1068

## RESULT 2

```

US-09-298-886-7
; Sequence 7, Application US/09298865
; Patent No. 6329170
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/298,886
; CURRENT FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1143)
; US-09-298-886-7

```





```
QY 670 GTGTTACAGCAACGGTATGCTGGTGGGAGAAATTAATGCTTCCCGAGGAGAC 729
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 780 GTGGTCAACGCGATGACATGGCTGGTGGGAGAGCATACAGCTCCCTTGGGGAC 839
QY 730 GTGGTGTCCGGGGAATGATATGAGGGTCCCGAGCCAAAGACTTCGGGCTGCTCAAC 789
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 840 GTGGTGTCCGCTGCAATGCCCTCCAGGGCTCACCTGCCAAGACTTCGCACTGCTCACA 899
QY 790 CAGTGAACACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCTCGGCA 849
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 900 CAGTGAACACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCTCGGCA 959
QY 850 GGTGGTATACCATCTACTTAGCAACTACACCTTCCCGGATTCCTCGTTCCCTCAAAGTC 909
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 960 GGGGGGACACTGCTACCTTGGGCACTTACCTTCCCACTCCCTTTCAAGTGGTC 1019
QY 910 TTAAACCCAGAGGCGCTTCTTACCCGAATGGGTGGGCAATCCCTGCCGATCT 962
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1020 TTAAACCCAGAGGCGCTTCTTACCCGAATGGGTGGGCAATCCCTGCCGATCT 1072

RESULT 6
US-09-151-592-1
; Sequence 1, Application US/09151592
; Patent No. 6355859
; GENERAL INFORMATION:
; APPLICANT: Bosworth, Brad
; APPLICANT: Ridpath, Julia
; APPLICANT: Wiseman, Barry
; TITLE OF INVENTION: INTERACTIONS BETWEEN GENOTYPE AND DIET IN SWINE THAT
; FILE REFERENCE: PREVENT E. COLI ASSOCIATED INTESTINAL DISEASE
; CURRENT APPLICATION NUMBER: US/09-151-592
; CURRENT FILING DATE: 1999-09-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Swine
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (9)..(1103)
; FEATURE:
; OTHER INFORMATION: FULL
US-09-151-592-1

Query Match 39.7%; Score 423.8; DB 4; Length 1269;
Best Local Similarity 69.6%; Pred. No. 2.2e-115;
Matches 591; Conservative 0; Mismatches 252; Indels 6; Gaps 1;

QY 136 GGCATGTTACAGATCAATTCATTGCCGCTGGGAAACAGATGGCGAATACGCCACA 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 GGCATGTTACAGATCAATTCATTGCCGCTGGGAAACAGATGGCGAATACGCCACA 302
QY 196 CTCCTTTGCACTGCCAGATGAAGCAAGCGCTTGGCTTCATCCCGCATCCAGCAAC 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 CTCCTGTCCTGGCGAGCTACAGCGCGCCAGCGCTTCATCCAGCTGCTATGCGCGC 362
QY 256 GCTCTAGGCGCCATCTTACAGATCAAGCTCCCGGTTTACACAGCGACAGCGCAAAAG 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 GTCCTGGCGCCCGTGTTCGCAATCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
QY 316 ATCCCATGGCAGATTAACATCTACAGCTGATGAGGAGGCTTACCGCCACATCCG 375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 GTCCTTGGCGGAGTGGAGCTTACAGCTGATGATGATGATGATGATGATGATGATGAT 482
QY 376 GGACACTTTGTCGCTTACGGGATACCGTGTCTCTGACCTTACACACACCTGCGG 435
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 GAGCCCTGGCTGAGCTACCGGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542
QY 436 CCAGAGATCCTGAAGGATTCACCTGCTATGACCAAGTGGGAGGAGGCCAGGCTTC 495
```

```
Db 543 GAGCAGATCGCAGGAGTTACCTTCACACACACCTTCGGCAAGAGGCCCGGGGTA 602
QY 496 CTGGGT-----GGTTCGGGGTGAATGGAGGACAGCCGAGTACTTTTGTGGGTGCAAT 549
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 603 CTGAGTCAGTTCGCTCTACCCCGCACAGGGGAGCGCCCGACACCTTCGTGGGGTGCAC 662
QY 550 GTGCGCGGAGGAGGATGTCATGTCATGCTTAATGTGTGGAGGGCGTGGTGGCTGAC 609
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 663 GTGCGCGCGGAGGATGTCATGTCATGCTTAATGTGTGGAGGGGCTGGTGGGTGAC 722
QY 610 CGSGGTTACCTGGAAAGAGCGCTTGGATATTCGGGCAAGCTATTTCATCTCAGTCTTC 669
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 723 GCGCTTACCTCCAGCAGGCTATGGACTGGTTCGGGCGCGATAGGAAGCCCGCTCTT 782
QY 670 GTGGTACAGCAACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 729
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 783 GTGGTCAACGACCAAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 842
QY 730 GTGGTGTTCGCGGCAATGCTATTGAGGGGTGCGCAAGGAGCTTCGCGCTGCTCACC 789
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 843 GTGATCTTGTGGGATGGCGGGAGCGCGCGCGCGAGGACITTCGCTGCTGGTG 902
QY 790 CAGTGAACACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCTACCTGCA 849
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 903 CAGTGAACACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCTACCTGCA 962
QY 850 GTGGTGTATACCATCTACTTAGCAACTACACCTTCCCGGATTCCTGCTTCCCTCAAAGTC 909
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 963 GGTGGAGATACCATCTACTTGGGTAACCTTACCTGCGGCTGCGGCTGCTGCTGCTGCT 1022
QY 910 TTAAAGCCAGAGGAGCGCTTCTTACCGGAATGGGTGGGCAATCCCTGCGGATCTGTCGCCA 969
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1023 TTAAAGCCAGAGGAGCGCTTCTTACCGGAATGGGTGGGCAATCCCTGCGGATCTGTCGCCA 1082
QY 970 CTCCTTAAAG 978
    ||||| |||
Db 1083 CTCAGATG 1091

RESULT 7
US-08-208-889A-1
; Sequence 1, Application US/08208889A
; Patent No. 5750176
; GENERAL INFORMATION:
; APPLICANT: Prieto, P., Smith, D., Cummings, R., Kopchick, J., Mukerji, P.,
; APPLICANT: Moremen, K., Pierce, J.
; TITLE OF INVENTION: Transgenic Production of Oligosaccharides and
; TITLE OF INVENTION: Glycoconjugates
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Donald O. Nickey
; ADDRESSEE: ROSS Products Division
; ADDRESSEE: Abbott Laboratories
; STREET: 625 Cleveland Avenue
; CITY: Columbus
; STATE: Ohio
; COUNTRY: United States of America
; ZIP: 43215
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS Version 6.21
; SOFTWARE: WordPerfect Version 6.0a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/208,889A
; FILING DATE: 09-MAR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA: No. 5750176 applicable
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614) 624-7080
; TELEFAX: (614) 624-3074
; TELEX: No. 5750176e
```

INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1155 base pairs  
 TYPE: Nucleic acid  
 STRANDEDNESS: Single  
 TOPOLOGY: Unknown  
 MOLECULE TYPE: Cloned cDNA representing the product of a  
 MOLECULE TYPE: human genomic DNA segment  
 DESCRIPTION: GDP-L-fucose- D-galactoside 2-alpha-fucosyl-transferase  
 HYPOTHETICAL:  
 ANTI-SENSE:  
 FRAGMENT TYPE: Entire amino acid sequence provided.  
 ORIGINAL SOURCE: Human Epidermal Carcinoma Cell line  
 ORGANISM:  
 STRAIN:  
 INDIVIDUAL ISOLATE:  
 DEVELOPMENTAL STAGE:  
 HAPLOTYPE:  
 TISSUE TYPE:  
 CELL TYPE:  
 CELL LINE:  
 ORGANELLE:  
 IMMEDIATE SOURCE: Human Epidermal Carcinoma Cell line  
 LIBRARY:  
 CLONE:  
 POSITION IN GENOME:  
 CHROMOSOME/SEGMENT: 19  
 MAP POSITION:  
 UNITS:  
 FEATURE:  
 NAME/KEY:  
 LOCATION:  
 IDENTIFICATION METHOD: DNA sequencing and restriction analysis  
 OTHER INFORMATION: The encoded product of nucleotide SEQ ID NO:  
 OTHER INFORMATION: 1: is the enzyme, GDP-L-fucose- D-galactoside 2-alpha-fucosyl-transferase  
 OTHER INFORMATION: which has the amino acid sequence described in SEQ. ID NO:1:  
 OTHER INFORMATION: is responsible for the synthesis of 2'-fucosyllactose.  
 PUBLICATION INFORMATION:  
 AUTHORS: V.P. Rajan, et al.  
 TITLE:  
 JOURNAL: Journal of Biological Chemistry  
 VOLUME: 264  
 ISSUE:  
 PAGES: 11158 - 11167  
 DATE: 1989  
 DOCUMENT NUMBER:  
 FILING DATE:  
 PUBLICATION DATE:  
 RELEVANT RESIDUES IN SEQ ID NO:  
 US-08-208-889A-1

Query Match 38.2%; Score 408.2; DB 1; Length 1155;  
 Best Local Similarity 68.7%; Pred. No. 8.5e-111;  
 Matches 579; Conservative 0; Mismatches 258; Indels 6; Gaps 1;

QY 136 GGCATGTCACGATCAATCCATTGGCGGCTGGGGAACACGATGGGCGGAATACGCCACA 195  
 DB 272 GGCACCTGGACTGTCTACCCCAATGGCGGTTGGTAAATCAGATGGGACAGTATGCCAG 331  
 QY 196 CTTCTTGACTGCCAGGATGACGACGGCTTGGCTTCATCCCGCATCCATGCACAC 255  
 DB 332 CTGCTGGCTCTGGCCAGCTCAACGGCGCGCGGCTTATCTGCTGCTCCATGATGCC 391  
 QY 256 GCTTAGCGCCCAATCTCAGGATAGCCTCCCGGTGTACACAGGACACGCGCCAAAAG 315  
 DB 392 GCCCGGCCCCGGTATTCGGATACCCCTCCCGGTGCTGGCGCCAGAGTGGACAGCCGC 451  
 QY 316 ATCCCATGCGAATTAACATCTCAACGACTGATGAGGAGCGTTACGCCACATTCGG 375  
 DB 452 ACGCCGTGGCGGAGTGCAGCTTCAGGACTGGATGCTCGGAGGAGTACGCGACTTGAGA 511  
 QY 376 GGACACTTTGCGGCTTACGGGATACCGGTGCTGCTGGACCTTACACCACTGGCC 435

DB 512 GATCCTTTCTCTGAAGCTCTCTGGCTTCCCTGCTCTTGGACTTCTTCCACCATCTCCGG 571  
 QY 436 CAGAGATCTCTGAAGAGTTCACCTGCTCATGACACAGTGGGGAGGAGGCCCTAGGCTTC 495  
 DB 572 GAACAGATCGCAGAGAGTTCACCTGCTCATGACACAGTGGGGAGGAGGCCCTAGGCTTC 631  
 QY 496 CTGCGTGGTCTCGGGGTG-----AATGGGAGCCAGCGAGTACTTTTGGGGTGGTCCAT 549  
 DB 632 CTGGGTCTAGCTCGCTCGGCTGGCCGCCACAGGGGACCGCCCGGCACCTTTGTCGGGCTCCAC 691  
 QY 550 GTGCGCGGAGGGGACTATGTGTCATGTCATGCTTAATGTGTGGAGGCGCTGGTGGTGGTGC 609  
 DB 692 GTGCGCGGAGGGGACTATGTGTCATGTCATGCTTAATGTGTGGAGGCGCTGGTGGTGGTGC 751  
 QY 610 CGGGGTACTTGGAAAGGCCCTGGATATGTTCCGGGCGAGCTTATTCATCTCCAGTCTTC 659  
 DB 752 AGCGCTTACTCTCGGCGAGGCCATGAGTGGTTCGGGCGACGACGAAAGCCCGCTTTC 811  
 QY 670 GTGGTTACAAGCAACGGTATGCTGCTGGCGGGAGAACATTAATGCTTCCCGAGGAGAC 729  
 DB 812 GTGGTCAACGACGACGCGCATGGTGTGTAAAGAAACATCGACACCTCCCGAGGCGAT 871  
 QY 730 GTGGTCTCGGGGCAATGTATGTAGGGGTGCCAGCCAGCCAGGACTTCGGGTGCTTCAC 789  
 DB 872 GTGACGTTTGTGGCGATGGACAGGAGGCTACACCGTGGAAAGACTTGGCCCTGCTCACA 931  
 QY 790 CAGTGCACACACACCATCATGACTATTTGGACCTTTGGGATTTGGGCTGCTACCTGGCA 849  
 DB 932 CAGTGCACACACACCATCATGACTATTTGGACCTTTGGGATTTGGGCTGCTACCTGGCT 991  
 QY 850 GGTGTGATACCATCTACTTAGCCAACTACACCTTCCGGATTCCTCGTTCCTCAAGTC 909  
 DB 992 GCGGAGAGACATCTCTACCTGGCCAACTTCCAGCTGCGACACTCTGAGTCTCGAAGATC 1051  
 QY 910 TTTAGCCAGAGACACCTTCTTACCCGATGGTGGGATCCCTGCCGATGTGTCCTCCA 969  
 DB 1052 TTTAGCCAGAGAGCGGCTTCTTCCCGATGGTGGGATTAATGACAGACTTGTCTCCA 1111  
 QY 970 CTC 972  
 DB 1112 CTC 1114

RESULT 8  
 US-08-433-271-1  
 ; Sequence 1, Application US/08433271  
 ; Patent No. 5891698  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Prieto, P., Smith, D., Cummings, R., Kopchick, J., Mukerji, P.,  
 ; APPLICANT: Moremen, K., Pierce, J.  
 ; TITLE OF INVENTION: Humanized Milk  
 ; NUMBER OF SEQUENCES: 1  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Donald O. Nicky  
 ; ADDRESSEE: ROSS Products Division  
 ; ADDRESSEE: Abbott Laboratories  
 ; STREET: 625 Cleveland Avenue  
 ; CITY: Columbus  
 ; STATE: Ohio  
 ; COUNTRY: United States of America  
 ; ZIP: 43215  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: MS-DOS Version 6.21  
 ; SOFTWARE: WordPerfect Version 6.0a  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/433,271  
 ; FILING DATE: 02-MAY-1995  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/209,122  
 ; FILING DATE: 09-MAR-1994

## CLASSIFICATION:

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (614) 624-7080

TELEFAX: (614) 624-3074

TELEX: No. 5891698e

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1155 base pairs

TYPE: Nucleic acid

STRANDEDNESS: Single

TOPOLOGY: Unknown

MOLECULE TYPE: Cloned cDNA representing the product of a

MOLECULE TYPE: human genomic DNA segment

DESCRIPTION: GDP-L-fucose- D-galactoside 2'-alpha-fucosyl-transferase

## HYPOTHETICAL:

## ANTI-SENSE:

FRAGMENT TYPE: Entire amino acid sequence provided.

ORIGINAL SOURCE: Human Epidermal Carcinoma Cell line

## ORGANISM:

## STRAIN:

INDIVIDUAL ISOLATE:

DEVELOPMENTAL STAGE:

HAPLOTYPE:

TISSUE TYPE:

CELL TYPE:

CELL LINE:

ORGANELLE:

IMMEDIATE SOURCE: Human Epidermal Carcinoma Cell line

## LIBRARY:

## CLONE:

POSITION IN GENOME:

CHROMOSOME/SEGMENT: 19

MAP POSITION:

UNITS:

## FEATURE:

## NAME/KEY:

## LOCATION:

IDENTIFICATION METHOD: DNA sequencing and restriction analysis

OTHER INFORMATION: The encoded product of nucleotide SEQ ID NO:

OTHER INFORMATION: 1: is the enzyme, GDP-L-fucose- D-galactoside 2'-alpha-fucosyl-transferase

OTHER INFORMATION: which has the amino acid sequence described in SEQ. ID NO:1:

OTHER INFORMATION: is responsible for the synthesis of 2'-fucosyllactose.

## PUBLICATION INFORMATION:

AUTHORS: V.P. Rajan, et al.

## TITLE:

JOURNAL: Journal of Biological Chemistry

VOLUME: 264

ISSUE:

PAGES: 11158 - 11167

DATE: 1989

DOCUMENT NUMBER:

FILING DATE:

PUBLICATION DATE:

RELEVANT RESIDUES IN SEQ ID NO:

US-08-433-271-1

Query Match

Best Local Similarity 38.2%; Score 408.2; DB 2; Length 1155;

Matches 579; Conservative 0; Mismatches 258; Indels 6; Gaps 1;

QY

136 GGCATGTTCAAGTCAATTCATTTGGCGGCTGGGGAACAGATGGGGAATACGCCACA 195

Db

272 GGCACCTGGAGCTGTCTACCCCAATGGCGGTTTGGTAATCAGATGGGAGATATGCCAAG 331

QY

196 CTTCTTGATGGCGGAGGATGACGAGCGCTGGCTTCATCCCGGCATCCATGGACACAC 255

Db

332 CTGCTGGCTCTGGCCAGCTCAACGCCGCCGGGCTTTATCTGCTGCCATGATGCC 391

QY

256 GCTCTAGCCGCCATCTTCAGGATACGCTCCCGGTGTACACAGGACGACGCCAAAAG 315

Db

392 GCCCTGGCCCCGGTATTCGGATCAACCTGCCGCTGGCCGCCAGAGTGGACAGCCGC 451

QY

316 ATCCCATGGCAGATTACCATCTCAACGACTGGATGGAGGACGGTACCGCCACATTCGG 375

452 ACCCGTGGGGGAGCTGCGAGCTTACGACTGATGCGGAGGAGTACGGGACTGAGA 511

376 GGACACTTTGTGGGCTTCAGGATACCCCTGCTCTGGACCTTCTACCAACCTGCGC 435

512 GATCCTTTCTGAGAGCTCTGCTGCTTCCCTGCTCTGGACTTCTTCCACCATCTCCG 571

436 CCAGAGATCTGAGGAGTTCACCTTGCATACCAAGTGGGGAGGAGGCCAGGCTTC 495

572 GAACAGATCGGAGAGTTCACCTTGCACGACCACTTCGGAAGAGGCGGAGAGTGTG 631

496 CTGCTGGTCTGCGGGTG-----AATGGAGCCAGCCGAGTACTTTGTGGGTGTCAT 549

632 CTGGGTCACTCGCCCTGGCCGCCACAGGGGACGCGCCGCGACCTTTGTGCGGCTCCAC 691

550 GTGCGCGGAGGAGTATGTGATGATGATGATGATGATGATGATGATGATGATGATG 609

692 GTGCGCGGAGGAGTATGTGATGATGATGATGATGATGATGATGATGATGATGATG 751

610 CGGGGTACCTGGAAAGGCCCTGGATATGTTCCGGGACAGCTTATTCATCTCCAGTCTC 669

752 AGCGCTACCTCGCGAGGCCATGGATGTTCCGGGACAGGCGGACGAAAGCCCGTTTC 811

670 GTGTTTACAAGCAACGGTATGGCTTGTGCTGGGAGAGCAATTAATGTTCCCGAGAGAC 729

812 GTGTTTACCAGCAACGGTATGGCTTGTGCTGGGAGAGCAATTAATGTTCCCGAGAGAC 871

730 GTGTTTCCCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 789

872 GTGACGTTTGTGGGATGGACAGAGGCTTACACCGTGGAAAGACTTTCGCTTCTCACA 931

790 CAGTGAACACACCATCATGATGATGATGATGATGATGATGATGATGATGATGATG 849

932 CAGTGAACACACCATCATGATGATGATGATGATGATGATGATGATGATGATGATG 991

850 GTGTTGATACCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 909

992 GCGGAGACATGCTTACCTGGCACTTCACTCCAGACTTCACTCCAGACTTCACTCCAG 1051

910 TTAAAGCCAGAGCAGCTTCCCTCCAGGATGGGAGGATGGGATGGGATGGGATGGGAT 969

1052 TTAAAGCCAGAGCAGCTTCCCTCCAGGATGGGAGGATGGGATGGGATGGGATGGGAT 1111

970 CTC 972

1112 CTC 1114

## RESULT 9

PCT-US91-00899-10

Sequence 10, Application PC/TUS9100899

GENERAL INFORMATION:

APPLICANT: Lowe, John B.

TITLE OF INVENTION: Method and Products For the Synthesis of

TITLE OF INVENTION: Oligosaccharide Structures on Glycoproteins, Glycolipids,

TITLE OF INVENTION: or as Free Molecules, and For the Isolation of Cloned

TITLE OF INVENTION: Genetic Sequences That Determine These Structures

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER &amp; NEUSTADT.

ADDRESS: P.C.

STREET: 1755 Jefferson Davis Highway, Suite 400

CITY: Arlington

STATE: Virginia

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/00899

FILING DATE: 19910214



```
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye Ph.D., Jean-Paul
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-021-55 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-5940
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 999 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Blood
PCT-US91-00899-10

Query Match 37.9%; Score 405; DB 5; Length 999;
Best Local Similarity 68.4%; Pred. No. 7e-110;
Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;

QY 136 GGCATGTTTCCACCAATTCATTGCGCGGTGGGAGACCAAGATGGGATACGCCACA 195
DB 136 GGCATGTTTCCACCAATTCATTGCGCGGTGGGAGACCAAGATGGGATACGCCACA 195
QY 139 GGCACCTGGAGTGTCTACCCCAATGCGCGGTGGGATACAGATGAGATGATGATG 198
DB 139 GGCACCTGGAGTGTCTACCCCAATGCGCGGTGGGATACAGATGAGATGATGATG 198
QY 136 CTCCTTGACCTGGCAGGATGACGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTG 255
DB 136 CTCCTTGACCTGGCAGGATGACGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTG 255
QY 139 CTGCTGGCTGTGGCCAGCTCAACGGCGCGCGCTTATCTCTCCCTCCCATGATGCC 258
DB 139 CTGCTGGCTGTGGCCAGCTCAACGGCGCGCGCTTATCTCTCCCTCCCATGATGCC 258
QY 256 GCTCTAGCGCCCACTTTCAGATGACCTGCGCGGTGTTTACAGGAGCACGGCCAAAAG 315
DB 256 GCTCTAGCGCCCACTTTCAGATGACCTGCGCGGTGTTTACAGGAGCACGGCCAAAAG 315
QY 259 GCGCTGGCGCCCGGATTCGCGATCACCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 318
DB 259 GCGCTGGCGCCCGGATTCGCGATCACCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 318
QY 316 ATCCCATGGCAGAAATACCACTCAACGACTGAGGAGGAGCGGTTCACCCGCAATCCGCGC 375
DB 316 ATCCCATGGCAGAAATACCACTCAACGACTGAGGAGGAGCGGTTCACCCGCAATCCGCGC 375
QY 319 AGCGCTGGCGGAGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 378
DB 319 AGCGCTGGCGGAGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 378
QY 379 GATCCTTTCCTGAAGCTCTCTGGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 438
DB 379 GATCCTTTCCTGAAGCTCTCTGGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 438
QY 436 CCAGAGATCCTGAAGAGTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 495
DB 436 CCAGAGATCCTGAAGAGTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 495
QY 439 CAACAGATCCCGCAGAGTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 498
DB 439 CAACAGATCCCGCAGAGTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 498
QY 496 CTGGGTGCTGCTGCGGGTG-----AATGGGAGCCAGCGAGTACTTTTGGGTGCTGCTG 549
DB 496 CTGGGTGCTGCTGCGGGTG-----AATGGGAGCCAGCGAGTACTTTTGGGTGCTGCTG 549
QY 499 CTGGGTGCTGCTGCGGGTG-----AATGGGAGCCAGCGAGTACTTTTGGGTGCTGCTG 558
DB 499 CTGGGTGCTGCTGCGGGTG-----AATGGGAGCCAGCGAGTACTTTTGGGTGCTGCTG 558
QY 550 GTGGCGCGAGGGAGTATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 609
DB 550 GTGGCGCGAGGGAGTATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 609
QY 559 GTGGCGCGTGGGAGTATCTCAGGTTATGCTCAGGCTGGAAGGTTGTTGGTGGCGAC 618
DB 559 GTGGCGCGTGGGAGTATCTCAGGTTATGCTCAGGCTGGAAGGTTGTTGGTGGCGAC 618
QY 610 CGGGGTACCTGGAAAGGCCCTGATGTTTGGGCGACGCTATTCATCTCCAGTCTTC 669
DB 610 CGGGGTACCTGGAAAGGCCCTGATGTTTGGGCGACGCTATTCATCTCCAGTCTTC 669
QY 619 AGCGCTTACCTCCGCGAGGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 678
DB 619 AGCGCTTACCTCCGCGAGGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 678
QY 670 GTGGTTACAAGCAAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 729
DB 670 GTGGTTACAAGCAAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 729
QY 679 GTGGTACCAACAGCGGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 738
DB 679 GTGGTACCAACAGCGGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 738
QY 730 GTGGTGTTCGGGCAATGTTATGAGGGTTCGCCAGCAAGGACTTCGCGCTGCTCACC 789
DB 730 GTGGTGTTCGGGCAATGTTATGAGGGTTCGCCAGCAAGGACTTCGCGCTGCTCACC 789
QY 739 GTGACGTTTGTGGGATGGAGAGGCTACACCTGGAAGACTTTCGCTGCTCACA 798
DB 739 GTGACGTTTGTGGGATGGAGAGGCTACACCTGGAAGACTTTCGCTGCTCACA 798
QY 790 CAGTGCAACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGGCTACCTGGCA 849
DB 790 CAGTGCAACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGGCTACCTGGCA 849
```

```
DB 799 CAGTGCACACACACCACTTATGACCATTTGGCACTTGGGCTTCTGGGCTGCCTACCTGGCT 858
QY 850 GGTGGTGATACCACTTACCTTACCACTTACCACTTACCACTTACCACTTACCACTTACCACT 909
DB 859 GCGGAGACACTGTCTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 918
QY 910 TTATAGCCAGAGCAGCCCTTCTACCCGATGGTGGGATCCCTCCGATGTCCTCCCA 969
DB 919 TTATAGCCGAGGCGGCTTCTGCGCGAGTGGTGGGATTAATGACACTTGTCTCCA 978
QY 970 CTC 972
DB 979 CTC 981

RESULT 10
US-08-395-800A-9
Sequence 9, Application US/08395800A
Patent No. 5807732
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B
APPLICANT: LENNON, GREGORY
APPLICANT: ROQUIER, SYLVIE
APPLICANT: GIORGI, DOMINIQUE
APPLICANT: KELLY, ROBERT J
TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE
TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES.DNA SEQUENCES ENCODING THE
TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSER: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,800A
FILING DATE: 28-FEB-1995
CLASSIFICATION: 435
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1136 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 39..1133
US-08-395-800A-9

Query Match 37.9%; Score 405; DB 1; Length 1136;
Best Local Similarity 68.4%; Pred. No. 7.4e-110;
Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;

QY 136 GGCATGTTTCCACCAATTCATTGCGCGGTGGGAGACCAAGATGGGATACGCCACA 195
DB 136 GGCATGTTTCCACCAATTCATTGCGCGGTGGGAGACCAAGATGGGATACGCCACA 195
QY 273 GGCACCTGGAGTGTCTACCCCAATGCGCGGTGGTGTATGATGAGATGGACATATGCCACG 332
DB 273 GGCACCTGGAGTGTCTACCCCAATGCGCGGTGGTGTATGATGAGATGGACATATGCCACG 332
QY 196 CTCCTTTCAGTGGCAGGATGATGAGGAGGCTTTCGCTTCTATCCCGCAATCATGCAAC 255
DB 196 CTCCTTTCAGTGGCAGGATGATGAGGAGGCTTTCGCTTCTATCCCGCAATCATGCAAC 255
QY 333 CTGCTGCTCTGGCCCACTCAACGCGCGGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTG 392
DB 333 CTGCTGCTCTGGCCCACTCAACGCGCGGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTG 392
```

256	QY	GCCTAGCGCCATCTTTCAGATCAGCCTCCCGGTGTTACACACGACACGGCCAAAAG	315
393	Db	GCCTGCGCCGGTATTTCGCATCAOCTTCCCGTGTGCCCCAGAAAGTGGACAGCGC	452
316	QY	ATCCATGGCAGATATACCATCTCAAGCACTGGATGGAGGAGCGTTACCGCCACATTCG	375
453	Db	ACGCGTGGCGGAGCTGCAGCTTCAGACTGTGATGTTCGAGGAGTAGCGGGACTTGAGA	512
376	QY	GGACACTTTGTGCGCTTCACGGATACCGTGTCTCTGGACCTTCTACCAACCTCGCG	435
513	Db	GATCCTTCTGAGCTCTTGCTTGCCCTGTCTCTGGACTTCTTCCACCACTCTCCGG	572
436	QY	CCAGAGATCTGAAGGAGTTACCCCTGCATGACACAGTGGGGAGGAGGCCAGCGCTTC	495
573	Db	GAACAGATCGCAGAGAGTTTCACTTGCACGACCACTTCGCGAAGAGGCGCAGTGTG	632
496	QY	CTGCGTGGCTGGGGTG-----AATGGAGGCCAGCCGAGTACTTTTGGGGTGTCCAT	549
633	Db	CTGGGTGAGTCTGCGCTTGGCCGACAGGGGACCGCCGCGCACTTGTGCGCGTCCAC	692
550	QY	GTGCGCGGAGGGACATGTGCATGTCACTGCTAAATGTGTGGAGGGCGTGTGGCTGAC	609
693	Db	GTGCGCGTGGGACATCTGCAGTTATCCCTCAGGCTTGGAGGGGTGTGGTGGCGGAC	752
610	QY	CGGGTTTACCTGAAAGGCCCTCGGATATGTTCCGGGACAGCTATTCACTCCAGTCTTC	659
753	Db	AGCGCTTACCTCGGAGGCCATGGACTGTTCCGGCACGGCACGAAGCCCGGTTTC	812
670	QY	GTGGTTTACAAGCAACGGTATGGCTGTGCGCGGAGAACATTAATGCTTCCGAGGAGAC	729
813	Db	GTGCTCACCAGCAACGGCATGAGTGGTGTAAAGAAACATCGACACTCCCAAGGCGAT	872
730	QY	GTGCTGTTCGCGGGCAATGGTATTGAGGGTCCGCCAGCCAGGACITCGCGCTGCTCAOC	789
873	Db	GTCAAGTTTCTGGCGATGGACAGGAGGCTACACCGTGGAAAGACTTTGGCCCTGCTACA	932
790	QY	CAGTGAACCAACACCATCATGACTATTGGACCTTTGGATTTGGGTGGCTACCTGGCA	849
933	Db	CAGTGAACCAACACCATTTATGACCATTTGCAOCTTCGGCTTTGGGTGCTACCTGGCT	992
850	QY	GGTGGTGTATACCATCTACTTAGCCACTACACCTTCCGGATTTCCGGTTCCTCAAGTC	909
993	Db	GGCGGAGACACTGTCTACCTGGCAACTTCAOCTTCCAGACTCTGAGTTCCTGAAGATC	1052
910	QY	TTTAAAGCAGAGGACGCTTCCTACCGAATGGTGGGCATCCCTCCGATCTGTCCCCA	969
1053	Db	TTTAAAGCGGAGCGGCCCTTCTTCCCGAGTGGTGGGCATTAATGACACTGTCTCCCA	1112
970	QY	CTC	972
1113	Db	CTC	1115

## RESULT 11

PCT-US95-07554-3

; Sequence 3, Application PC/TUS9507554

GENERAL INFORMATION:

APPLICANT: Sandrin, Mauro S.

; APPLICANT: Fodor, William L.

; APPLICANT: Rother, Russell P.

APPLICANT: Squinto, Stephen P.

APPLICANT: McKenzie, Ian F. C.

;	TITLE OF INVENTION:	Methods for Reducing Hypersensitive Rejection
;	TITLE OF INVENTION:	

1 TITLE OF INVENTION: Hyperacute Rejection of Xenografts  
2  
3 NUMBER OF SEQUENCES: 5

NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Maurice M. Klee

STREET: 1951 Burr Street

; CITY: Fairfield

; STATE: Connecticut

; COUNTRY: USA  
; ZTD: 06420; ZIP: 06430  
; COMPUTER READ

; COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 750 kb storage  
COMPUTER: Dell 486/50  
OPERATING SYSTEM: DOS 6.2  
SOFTWARE: Wordperfect 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/07554  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/260,201  
FILING DATE: June 15, 1994  
CLASSIFICATION:  
APPLICATION NUMBER: 08/278,282  
FILING DATE: July 21, 1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Klee, Maurice M.  
REGISTRATION NUMBER: 30,399  
REFERENCE/DOCKET NUMBER: ALX-144.1PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 255-1400  
TELEFAX: (203) 254-1101  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1174 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Double  
TOPOLOGY: Linear  
MOLECULE TYPE: cDNA to mRNA  
DESCRIPTION: Human H-transferase  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
PUBLICATION INFORMATION:  
AUTHORS: Larsen, R.D.  
AUTHORS: Ernst, L.K.  
AUTHORS: Nair, J.B.  
AUTHORS: Lowe, J.R.P.  
TITLE: Molecular cloning, sequence, and  
expression of a human Gp-L-fucosyl-  
D-galactoside 2-alpha-L-  
fucosyltransferase cDNA that encodes  
form the H blood group antigen  
JOURNAL: Proceedings of the National  
Academy of Sciences, USA  
VOLUME: 87  
PAGES: 6674-6678  
DATE: SEP-1990  
PCT-US95-07554-3

Query Match 37.9% Score 405 DB 5: Length 1174:

Best Local Similarity 68.4%; pred. No. 7.5e-110;

Matches	577;	Conservative	0;	Mismatches	260;	Indels	6;	Gaps	1;
---------	------	--------------	----	------------	------	--------	----	------	----

QY 136 GGCATGTTACGATCAATTCCATTGGCCGGCTGGGGAACCAGATGGCGGAATACGCCACA 195

[illegible]

Db 249 GGCACCTGGACTGTCTACCCCAATGGCCGGTTGGTAATCAGATGGGACAGTAIGCCACG 308

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

196 CTCTTTGCACTGGCCAGGATGAACGGACGGCNTGCGTTTCATCCCCGCATCCATGCACAAC 255

db 309 CTGCTGGCTTGGCCACCTCAACGGCCCGCCCTTATCTCCATCTCTCC 360

DB 309 CTGCTGGCTCTGGCCGAGCTCAACGGCCGCCGGGCGCTTTATCCTGCCATGCATGCC 368

QY 256 GCTCTAGCGCCCATCTTCAGGATCAGCCTCCCGGTGTTACACAGCGACGGCCAAAAG 315

[illegible]

Ddb 369 GCCCTGGCCCCGGTATTCCGCATCACCCCTGCCCTGCTGGCCCCCAGAAAGTGACAGCCGC 428



```

QY 970 CTC 972
Db 1125 CTC 1127

RESULT 13
US-08-675-773B-4
; Sequence 4, Application US/08675773B
; Patent No. 6166288
; GENERAL INFORMATION:
; APPLICANT: DIAMOND, LISA E
; APPLICANT: LOGAN, JOHN S
; APPLICANT: BYRNE, GUERARD W
; APPLICANT: SHARMA, AJAY
; TITLE OF INVENTION: METHOD OF PRODUCING TRANSGENIC ANIMALS
; TITLE OF INVENTION: FOR XENOTRANSPLANTATION.
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESS: BROWDY AND NEIMARK, P.L.L.C.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,773B
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, IVER P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: DIAMOND=1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2268 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-675-773B-4

Query Match 37.9%; Score 405; DB 4; Length 2268;
Best Local Similarity 68.4%; Pred. No. 1e-109;
Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;

QY 136 GGCAGTTTCACGATCAATTCATTCGCGGCTGGGGAACAGATGGGGAATACGCCACA 195
Db 749 GGCACCTGGAGCTCTACCCCAATGCGCGTTTGTATATCAGATGGACAGATGCCAG 808

QY 196 CTTTGTGCTGGCAGGATGACGACGCGCTTGCCTTATCCCGCATCCATGACACAC 255
Db 809 CTGCTGGCTTGGCGGAGCTACGACGCGCGCGCGCTTATCTGCTGCCATGCTGCC 868

QY 256 GTCTAGCGCCCATCTTCAGATCACTTCAGACTGATGAGGAGCGTTTACCGCCACATTCGG 315
Db 869 GCCCTGGCGCGGATTCGCGCATCACCCTGCGCGCTGCTGCGCCAGAGATGACAGCCGC 928

QY 316 ATCCCATGACAGATTTACCATCTTCACGACTGATGAGGAGCGTTTACCGCCACATTCGG 375
Db 929 ACGCCGTGGCGGAGCTGACGCTTCACGACTGATGCTGAGGAGTACCGGACITGAGA 988

QY 376 GGACACTTTGTGGCTTACCGGATACCGGCTCTCTGGACCTTCTACACACCTGCGC 435
Db 989 GATCCTTCTGAAAGCTCTGCTGCTTCCCTGCTTGGACTTCTTCCACCATCTCCGG 1048

```

```

RESULT 14
US-08-273-411-2
; Sequence 2, Application US/08273411
; Patent No. 5625124
; GENERAL INFORMATION:
; APPLICANT: Falk, Jeffrey I.
; APPLICANT: Gordon, Jeffrey I.
; TITLE OF INVENTION: Animal Model for Gastro-Intestinal
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,411
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU106
; TELECOMMUNICATION INFORMATION:

```

[illegible]

```

;
; AUTHOR: Larsen, et al.
; JOURNAL: Proc. Nat'l Acad. Sci. USA
; VOLUME: 87
; PAGES: 6674-6678
; DATE: 1990
;
; RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 3373
; US-08-273-411-2

Query Match          37.9%; Score 405; DB 1; Length 3373;
Best Local Similarity 68.4%; Pred. No. 1.2e-109;
Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;

QY 136 GGCATGTCACGATCAATTCATTGGCGCGCTGGGACACAGATGGCGGAATACGCCACA 195
Db 338 GGCACCTGGACTCTCTACCCCAATGGCGGGTTGGTAATCAGATGGGACAGATATGCCACG 397
QY 196 CTCTTTGCATCGCCAGGAGTGAACGGACGGGCTTGGCTTCATCCCGCGCATCGATGCACAAAC 255
Db 398 CTGCTGGCTCTGGCCCAAGCTTCACGGCGCGCGGCGCTTATCTCTCCCTGCCATGCCATGCC 457
QY 256 GCTCTAGCGCCCATCTTTCAGATCAGCGCTCCCGGTGTTTACAGCGACACGCGCCAAAAG 315
Db 458 GCCTGTGCCCCGGTATTCCGCATCACCTGTCCGCTGTGGCCCCAGAAAGTGACACGCCGC 517
QY 316 ATCCCATGGCAGAAATACCACTCTCAAGCACTGGATGGAGGAGCGTACCGCCACATTCGG 375
Db 518 ACCCGGTGGCGGAGCTGCAGCTTCACGACTGGATGTCGGAGGAGTAGCGGACATTGAGA 577
QY 376 GGACACTTTGTGCGCTTCACGGGATACCGGTGCTCCTGGACCTTCTACACCACTGGCG 435
Db 578 GATCCTTTCCTGAAGCTCTCTGGGCTTCCCGTGCTTTGGACTTTCTCCACCATCTCCGG 637
QY 436 CCAGAGATCTGANGAGATTACCCCTGCATGACACAGCTGCGGGGAGGAGCGCCAGCCCTTC 495
Db 638 GAACAGATCGCAGAGAGATTACCTCGACACACCCTTCGCGGAGAGGCGCAGAGTGTG 697
QY 496 CTGCGTGGTCTGCGGGTG-----AATGGGAGCGACCGGAGTACTTTTGTGGGTGTCCAT 549
Db 698 CTGGTGCAGCTCCGCTTGGGCGGCACAGGGGAGCCGCCCGCCACCTTTGTGCGCGTCCAC 757
QY 550 GTGCGCCGAGGGACTATGTCATGTGATGTGCTTAATGTGTGGAAGGGGTGTGTGCTGCAC 609
Db 758 GTGCGCGGTGGGACTATCTGCAGTTATGCTCAGCGCTGGAAGGGTGTGTGTGGCGCAC 817

```

Qy	610	CGGGTTACCTGGAAAGGCCCTGGATATGTTCCGGGACGCTATTCATCTCCATCTTC	669
Db	818	AGCCCTACCTCCGGCAGGCCATGGACTGGTTCGGGACGGCAGAACCCCGGTTTC	877
Qy	670	GTGGTTACAAGCAACGGTATGSCCTGGTCCGGGAGACATTAATGCTTCCCGAGGAC	729
Db	878	GTGGTCACCAACACGGGCATGGAGTGGTGAAGAAACATCGACACCTCCCGAGGCGAT	937
Qy	730	GTGGTGTTCCGGGGCAATGGTATTGAGGGTGCACCGCAAGGACTTGCCTGTCTCAC	789
Db	938	GTGACGTTTGTGGCGATGGACGAGGCTACCCGTGGAAGACTTGGCCCTGCTACA	997
Qy	790	CAGTGCACCAACACCATCAIGACTTAATGGGACCTTTGGGATTTGGGCTGCTACCTGGCA	849

Db 2413 GGCCTGGCCCGGTAATTCGGCATCCCTGCGCGTCTGSCCCACGAAGTGGACACCGC 2472  
QY 316 ATCCCATGCGAGAAATACCATCTCAAGACTGGATGGAGAGGTTACCGCCACATTCGG 375  
Db 2473 AGCCCTGGCGGAGCTGAGCTTCAGACTGGATGTCGAGGAGTACGGGACTTGAGA 2532  
QY 376 GGACACTTTGCGGCTTACGGGATACCGGTCTGCTGCGACCTTCTACGACCACTCGGC 435  
Db 2533 GATCCCTTCTGAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2592  
QY 436 CCAGAGATCCTGAAGAGTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 495  
Db 2593 GAACAGATCGGAGAGAGTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2652  
QY 496 CTGCGTGGTCTGGGGTG-----AATGGAGGACCGGAGTACTTTTGGGTGTCCAT 549  
Db 2653 CTGGGTCAAGCTCGGCTGGGCGGCGACAGGGGACCGCGCGACCTTTGTCGGCGTCCAC 2712  
QY 550 GTGGCGCGAGGGACTATGTGCATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 609  
Db 2713 GTGGCGGCTGGGGACTATGTGCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2772  
QY 610 CGGGGTTACTGGAAGGCGCTGGATATGTCGGGCGAGCTATTCATCTCCAGTCTTC 669  
Db 2773 AGCGCTTACTTCGGCAGGCGATGGACTGTTCCGGGCGAGCGCAAGCGCCGTTTC 2832  
QY 670 GTGGTTACAGCAACGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 729  
Db 2833 GTGGTCACGAGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2892  
QY 730 GTGGTCTTCGGGCGCAATGCTATGAGGGGTGCGGCGAGGACTTGGGCTGCTGCTGCTG 789  
Db 2893 GTGACGTTTGTGGCGATGGAGAGGCTACACCGTGAAGACTTTGCCCTGCTGCTGCTG 2952  
QY 790 CAGTGAACCAACCATCATGCTATGAGGACTTTGGGACTTTGGGACTTTGGGACTTTGGG 849  
Db 2953 CAGTGAACCAACCATCATGCTATGAGGACTTTGGGACTTTGGGACTTTGGGACTTTGGG 3012  
QY 850 GTGGTGTATACCATCTACTAGCCACTACACCTTCGCGATTCCTCCGTTCTCCTCAAGTC 909  
Db 3013 GCGGAGACACTGTCTACCTGGGCGCACTTCACCTGCCAGACTCTGAGTTCTCTGAAGATC 3072  
QY 910 TTATAGCCAGAGCGACCTTCTACCGAATGGGTGGGATCCCTGCGGATCTGTCGCCCA 969  
Db 3073 TTATAGCCAGAGCGGCGCTTCTGCGGAGTGGGTGGGATTAATGACACTTGTCTCCA 3132  
QY 970 CTC 972  
Db 3133 CTC 3135

Search completed: May 26, 2003, 11:58:34  
Job time : 60.2896 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 11:58:37 ; Search time 163.789 Seconds  
(without alignments)  
8610.191 Million cell updates/sec

Title: US-10-040-863-9

Perfect score: 1068

Sequence: 1 CTCACGAGCGAATAGTGGAA.....GAAGAGCCTCTGATGGAA 1068

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_NA:\*

```
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PC1_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PC1US_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1068	100.0	1068	10	US-09-999-672-9
2	1068	100.0	1068	12	US-10-040-863-9
3	1068	100.0	1068	10	US-09-999-672-7
4	1068	100.0	1068	12	US-10-040-863-7
5	612.8	57.4	1043	10	US-09-051-034A-1
6	589.4	55.2	3088	10	US-09-954-456-45
7	589.4	55.2	3088	10	US-09-954-456-1821
8	589.4	55.2	3088	10	US-09-969-347-234
9	423.8	39.7	1098	10	US-09-051-034A-3
10	423.8	39.7	1269	10	US-09-844-268-12
11	423.8	39.7	1269	10	US-09-844-705-12
12	405	37.9	3373	9	US-10-105-963-9
13	405	37.9	8174	10	US-09-863-475A-5
14	72.8	6.8	100	10	US-09-999-672-29
15	72.8	6.8	100	12	US-10-040-863-29
16	71.2	6.7	100	10	US-09-999-672-23
17	71.2	6.7	100	10	US-09-999-672-26
18	71.2	6.7	100	12	US-10-040-863-23
19	71.2	6.7	100	12	US-10-040-863-26

20	69.6	6.5	100	10	US-09-999-672-20	Sequence 20, Appl	
21	69.6	6.5	100	12	US-10-040-863-20	Sequence 20, Appl	
22	63.2	5.9	100	10	US-09-999-672-14	Sequence 14, Appl	
23	63.2	5.9	100	10	US-09-999-672-17	Sequence 17, Appl	
24	63.2	5.9	100	12	US-10-040-863-14	Sequence 14, Appl	
25	63.2	5.9	100	12	US-10-040-863-17	Sequence 17, Appl	
c	26	40.4	3.8	3044	9	US-10-125-540-98	Sequence 98, Appl
c	27	40.4	3.8	3044	10	US-09-764-870-98	Sequence 98, Appl
	28	38.4	3.6	50	10	US-09-999-672-16	Sequence 16, Appl
	29	38.4	3.6	50	12	US-10-040-863-16	Sequence 16, Appl
	30	37.6	3.5	1569	9	US-10-051-643-113	Sequence 113, App
	31	37.6	3.5	1569	9	US-09-880-505-113	Sequence 113, App
	32	37.6	3.5	1626	9	US-10-051-643-159	Sequence 159, App
	33	37.6	3.5	1626	9	US-09-880-505-159	Sequence 159, App
	34	37.4	3.5	50	10	US-09-999-672-28	Sequence 28, Appl
	35	37.4	3.5	50	12	US-10-040-863-28	Sequence 28, Appl
	36	37	3.5	50	10	US-09-999-672-25	Sequence 25, Appl
	37	37	3.5	50	12	US-10-040-863-25	Sequence 25, Appl
	38	36.8	3.4	50	10	US-09-999-672-19	Sequence 19, Appl
	39	36.8	3.4	50	12	US-10-040-863-19	Sequence 19, Appl
	40	36.4	3.4	802	9	US-10-184-634-312	Sequence 312, App
	41	36.4	3.4	802	9	US-10-184-634-312	Sequence 312, App
c	42	36.4	3.4	2030	10	US-09-880-107-2416	Sequence 2416, Ap
	43	36.2	3.4	916	9	US-10-184-644-78	Sequence 78, Appl
	44	36.2	3.4	916	9	US-10-184-634-78	Sequence 78, Appl
	45	36	3.4	1416	8	US-08-900-220C-19	Sequence 19, Appl

## ALIGNMENTS

## RESULT 1

US-09-999-672-9

; Sequence 9, Application US/099999672

; Patent No. US20020127655A1

; GENERAL INFORMATION:

; APPLICANT: Eric H. Holmes et al.

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE

; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 8511-029

; CURRENT APPLICATION NUMBER: US/09/399,672

; PRIOR FILING DATE: 2001-10-31

; PRIOR APPLICATION NUMBER: US/09/298,886

; PRIOR FILING DATE: 1999-04-26

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 9

; LENGTH: 1068

; TYPE: DNA

; ORGANISM: Rattus norvegicus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1062)

; US-09-999-672-9

Query Match 100.0%; Score 1068; DB 10; Length 1068;

Best Local Similarity 100.0%; Pred. No. 6;

Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCACGAGCGAATAGTGGAAACACAGAACCCCTCTCAGAGAGATCGGCGACAGCGAGCATGGG 120

Db 1 CTCACGAGCGAATAGTGGAGCTCCACCCCTCTCAGAGAGATTCACGATGAGCACT 60

QY 61 CAATATCTCTCGGAAACACAGAACCCAGAGATCGGCGACAGCGAGCATGGG 120

Db 61 CAATATCTCTCGGAAACACAGAACCCAGAGATCGGCGACAGCGAGCATGGG 120

QY 121 AATGGAGACTCGGGGCATGTTTCAGATCAATTCATTGGCGGTGGGAACAGATG 180

Db 121 AATGGAGACTCGGGGCATGTTTCAGATCAATTCATTGGCGGTGGGAACAGATG 180

QY 181 GGCGAATACGCCACACTCTTTCAGTGGCCAGGATGACGAGCGGTTCGGTTCATCCCC 240

181 GCGCAATACCCACACATCTTTGGCACTGGCCAGATGACCGAGCGCTTGGCTTCATCCCC 240  
241 GCATCATGACAAACGCTCTAGAGCCCATCTTCCAGATCAGCCTCCCGTGTGTACACAGC 300  
241 GCATCATGACAAACGCTCTAGAGCCCATCTTCCAGATCAGCCTCCCGTGTGTACACAGC 300  
301 GACAGGCCAAAAGATCCCATGCGAGAAATACCATCTCAACAGCTGGATGGAGGAGCGT 360  
301 GACAGGCCAAAAGATCCCATGCGAGAAATACCATCTCAACAGCTGGATGGAGGAGCGT 360  
361 TACCGCCACATTCGCGGACACTTTGTGCGCTTCAACGGATACCGGTCTCCTCGACCTTC 420  
361 TACCGCCACATTCGCGGACACTTTGTGCGCTTCAACGGATACCGGTCTCCTCGACCTTC 420  
421 TACACACCTGCGGCCAGAGATCTCTGAAGAGTTTCAACCTGCATGACCAACGCTGCGGAG 480  
421 TACACACCTGCGGCCAGAGATCTCTGAAGAGTTTCAACCTGCATGACCAACGCTGCGGAG 480  
481 GAGGCCAGGCTTCTCGCTGCTGCTGCGGTGAATGGAGCCAGCGAGTCTTTTGTG 540  
481 GAGGCCAGGCTTCTCGCTGCTGCTGCGGTGAATGGAGCCAGCGAGTCTTTTGTG 540  
541 GGTGTCATCTGCGCGAGGAGTATGTGATGTGATGCTATGCTGCTGCTGCTGCTGCTGCTG 600  
541 GGTGTCATCTGCGCGAGGAGTATGTGATGTGATGCTATGCTGCTGCTGCTGCTGCTGCTG 600  
601 GTGGCTGACGGGTTTACCTGGAAGGCGCTGATGATGTTGCGGCGACGCTATTCATCT 660  
601 GTGGCTGACGGGTTTACCTGGAAGGCGCTGATGATGTTGCGGCGACGCTATTCATCT 660  
661 CCAGTCTTCTGCTTACAAGCAACGCTATGGCTGCTGCGGAGCAACATTAATGCTTCC 720  
661 CCAGTCTTCTGCTTACAAGCAACGCTATGGCTGCTGCGGAGCAACATTAATGCTTCC 720  
721 CGAGGAGAGCTGGTGTTCGCGGCAATGATGATGAGGGTTCGCGAGCAAGGACTTCGCG 780  
721 CGAGGAGAGCTGGTGTTCGCGGCAATGATGATGAGGGTTCGCGAGCAAGGACTTCGCG 780  
781 CTGCTCAGCCAGTGCACACACACATCATGATGATGAGGGTTCGCGAGCAAGGACTTCGCG 840  
781 CTGCTCAGCCAGTGCACACACACATCATGATGATGAGGGTTCGCGAGCAAGGACTTCGCG 840  
841 TACCTGGCAGGTGGTATACCATCTTACTTAGCAACTACACCTTCCCGATTCCTCGTTC 900  
841 TACCTGGCAGGTGGTATACCATCTTACTTAGCAACTACACCTTCCCGATTCCTCGTTC 900  
901 CTCAGAGTCTTTAAGCCAGGAGCGCTTCTACCGCAATGGGTGGGCAATCCCTGCGGAT 960  
901 CTCAGAGTCTTTAAGCCAGGAGCGCTTCTACCGCAATGGGTGGGCAATCCCTGCGGAT 960  
961 CTGCTCCCACTCCTTAAGCATTAACACAGCGCTGCTGCGGCTCCCACTTCCACCTCAAG 1020  
961 CTGCTCCCACTCCTTAAGCATTAACACAGCGCTGCTGCGGCTCCCACTTCCACCTCAAG 1020  
1021 GCAAAAGAGTCACTTGTACCTGCGAGGAGCGCTTCTGATGGAA 1068  
1021 GCAAAAGAGTCACTTGTACCTGCGAGGAGCGCTTCTGATGGAA 1068

## RESULT 2

US-10-040-863-9  
; Sequence 9, Application US/10040863  
; Patent No. US20020137165A1  
; GENERAL INFORMATION:  
; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: 8511-029  
; CURRENT APPLICATION NUMBER: US/10/040,863  
; CURRENT FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: 09/298,886

; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 1068  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1062)  
US-10-040-863-9

Query Match 100.0%; Score 1068; DB 12; Length 1068;  
Best local Similarity 100.0%; Pred. No. 0;  
Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTCACGAGCGAATAGTGAAGCTCCACCCCTGTTCAGAGAAGGAATACCGATGACGACT 60  
DB 1 CTCACGAGCGAATAGTGAAGCTCCACCCCTGTTCAGAGAAGGAATACCGATGACGACT 60  
QY 61 CAATATGCTCTCGGGAACACAGAAAGCCAGAGATGCGAGCGGACACGAGCAGCATGGG 120  
DB 61 CAATATGCTCTCGGGAACACAGAAAGCCAGAGATGCGAGCGGACACGAGCAGCATGGG 120  
QY 121 AATGGAGAGCTGCGGGCATGTTACAGATCAATTCATTCGCGCGCTGGGGAACCATG 180  
DB 121 AATGGAGAGCTGCGGGCATGTTACAGATCAATTCATTCGCGCGCTGGGGAACCATG 180  
QY 181 GCGCAATACCCACACACTCTTTGCACTGGCCAGATGACGAGCGGTGGCTTCATCCCC 240  
DB 181 GCGCAATACCCACACTCTTTGCACTGGCCAGATGACGAGCGGTGGCTTCATCCCC 240  
QY 241 GCATCATGACAAACGCTCTAGAGCCCATCTTCCAGATCAGCCTCCCGTGTGTACACAGC 300  
DB 241 GCATCATGACAAACGCTCTAGAGCCCATCTTCCAGATCAGCCTCCCGTGTGTACACAGC 300  
QY 301 GACAGGCCAAAAGATCCCATGCGAGAAATACCATCTCAACAGCTGGATGGAGGAGCGT 360  
DB 301 GACAGGCCAAAAGATCCCATGCGAGAAATACCATCTCAACAGCTGGATGGAGGAGCGT 360  
QY 361 TACCGCCACATTCGCGGACACTTTGTGCGCTTCAACGGATACCGGTCTCCTCGACCTTC 420  
DB 361 TACCGCCACATTCGCGGACACTTTGTGCGCTTCAACGGATACCGGTCTCCTCGACCTTC 420  
QY 421 TACCGCCACCTGCGGCCAGAGATCCTGAAGAGTTTCAACCTGCATGACCAACGCTGCGGAG 480  
DB 421 TACCGCCACCTGCGGCCAGAGATCCTGAAGAGTTTCAACCTGCATGACCAACGCTGCGGAG 480  
QY 481 GAGGCCAGGCTTCTCGCTGCTGCTGCGGTGAATGGAGCCAGCGAGTCTTTTGTG 540  
DB 481 GAGGCCAGGCTTCTCGCTGCTGCTGCGGTGAATGGAGCCAGCGAGTCTTTTGTG 540  
QY 541 GGTGTCATCTGCGCGAGGAGTATGTGATGATGCTATGCTGCTGCTGCTGCTGCTGCTG 600  
DB 541 GGTGTCATCTGCGCGAGGAGTATGTGATGATGCTATGCTGCTGCTGCTGCTGCTGCTG 600  
QY 601 GTGGCTGACGGGTTTACCTGGAAGGCGCTGATGATGTTGCGGCGACGCTATTCATCT 660  
DB 601 GTGGCTGACGGGTTTACCTGGAAGGCGCTGATGATGTTGCGGCGACGCTATTCATCT 660  
QY 661 CCAGTCTTCTGCTTACAAGCAACGCTATGGCTGCTGCGGAGCAACATTAATGCTTCC 720  
DB 661 CCAGTCTTCTGCTTACAAGCAACGCTATGGCTGCTGCGGAGCAACATTAATGCTTCC 720  
QY 721 CGAGGAGAGCTGGTGTTCGCGGCAATGATGATGAGGGTTCGCGAGCAAGGACTTCGCG 780  
DB 721 CGAGGAGAGCTGGTGTTCGCGGCAATGATGATGAGGGTTCGCGAGCAAGGACTTCGCG 780  
QY 781 CTGCTCAGCCAGTGCACACACACATCATGATGATGAGGGTTCGCGAGCAAGGACTTCGCG 840  
DB 781 CTGCTCAGCCAGTGCACACACACATCATGATGATGAGGGTTCGCGAGCAAGGACTTCGCG 840  
QY 841 TACCTGGCAGGTGGTATACCATCTTACTTAGCAACTACACCTTCCCGATTCCTCGTTC 900



Db 301 LKVKPEAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 353

RESULT 3  
AAB36104

ID AAB36104 standard; Protein; 380 AA.

XX AC AAB36104;

XX DT 19-FEB-2001 (first entry)

XX DE Rat hepatoma H35 cell alpha1-2fucosyltransferase.

XX KW Rat; alpha1-2fucosyltransferase; cytostatic; neuroprotective;  
KW neotrophic; gene therapy; Fucalpal-2Galbeta1-3GalNAc; immunotherapy;  
KW immunosuppression; cancer; neurological disease;  
KW small cell lung carcinoma.

XX OS Rattus norvegicus.

XX PN WC200064464-A1.

XX PD 02-NOV-2000.

XX PF 23-APR-1999; 99WO-US07384.

XX PR 23-APR-1999; 99WO-US07384.

XX PA (PACI-) PACIFIC NORTHWEST CANCER FOUND.

XX PI Holmes EH, Sherwood AL;

XX WPI; 2000-687262/57.

XX N-PSDB; AAG67965.

XX New rat ganglioside GM1-specific alpha1-2Fucosyltransferase, useful for  
PT preparation of fucosyl GM1 which is useful as a nutritional composition  
PT or immunotherapeutic for cancer and neurological diseases -

XX Claim 1; Fig 5; 91pp; English.

XX The present sequence was given in a specification relating to an isolated  
CC rat ganglioside GM1-specific alpha1-2fucosyltransferase protein. The  
CC protein or its cellular fraction is useful for synthesis of a molecule  
CC comprising Fucalpal-2Galbeta1-3GalNAc, a glycolipid, glycoprotein,  
CC glycolipoprotein or a free oligosaccharide comprising  
CC Fucalpal-2Galbeta1-3GalNAc. The method involves contacting  
CC alpha1-2fucosyltransferase with GDP-fucose and a molecule or glycolipid,  
CC glycoprotein, glycolipoprotein or oligosaccharide having a terminal  
CC Galbeta1-3GalNAc group. It is also useful for synthesis of fucosyl-GM1  
CC by contacting the protein with GDP-fucose and ganglioside GM1. The  
CC obtained glycoproteins, glycolipoproteins, glycolipids and  
CC oligosaccharides are useful as nutritional compositions and fucosyl-GM1  
CC is useful for inducing an immunotherapeutic or immunosuppressive action  
CC against cancer, neurological disease or small cell lung carcinoma.

XX Sequence 380 AA;

Query Match 100.0%; Score 1896; DB 21; Length 380;  
Best Local Similarity 100.0%; Pred. No. 1.3e-198;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQQRIVKLQPLSEKELPMTTOMSGNTEPEMRDSDQHGNGELRGMTINSIGRLGNQM 60  
DB 28 LQQRIVKLQPLSEKELPMTTOMSGNTEPEMRDSDQHGNGELRGMTINSIGRLGNQM 87

QY 61 GEYATLALARMNGRLAFIPASHNNAAPFRISLPVLSHSDTAKKIPQWNYHLNDWMEER 120  
DB 88 GEYATLALARMNGRLAFIPASHNNAAPFRISLPVLSHSDTAKKIPQWNYHLNDWMEER 147

QY 121 YRHIPGHVFTGYPCSWTYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSQPSITFV 180  
DB 148 YRHIPGHVFTGYPCSWTYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSQPSITFV 207

QY 181 GVHVRGDIYHVMFNWKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCENINAS 240  
DB 208 GVHVRGDIYHVMFNWKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCENINAS 267

QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTFGWAAYLAGGDTIYLYLTPDSPF 300  
DB 268 RGDVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTFGWAAYLAGGDTIYLYLTPDSPF 327

QY 301 LKVKPEAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 353

DB 328 LKVKPEAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 380

RESULT 4  
AAE16622

ID AAE16622 standard; Protein; 380 AA.

XX AC AAE16622;

XX DT 09-APR-2002 (first entry)

XX DE Rat hepatoma H35 cell alpha1-2Fuct.

XX KW Rat; alpha1-2fucosyltransferase; alpha1-2Fuct; antisense therapy;  
KW galactose beta1-3N-acetylgalactosamine; Galbeta1-3GalNAc; glycolipid;  
KW glycoprotein; glycolipoprotein; oligosaccharide; fucosyl-GM1; cancer;  
KW gene therapy; oncogenic transformation; cytostatic; ganglioside; GM1;  
KW cell transformation.

XX OS Rattus norvegicus.

XX FH Key Location/Qualifiers

XX FT Domain 1..27 /note= "Intracellular/Transmembrane domain"

XX FT Modified-site 199 /note= "N-glycosylated"

XX FT Region 221..380

XX FT Modified-site 265 /note= "Region which overlaps rat FTE"

XX FT Modified-site 293 /note= "N-glycosylated"

XX FT Modified-site 293 /note= "N-glycosylated"

XX US6329170-B1.

XX 11-DEC-2001.

XX 23-APR-1999; 99US-0298886.

XX 23-APR-1999; 99US-0298886.

XX (NWHO-) NORTHWEST HOSPITAL.

XX Holmes EH, Sherwood AL;

XX WPI; 2002-121132/16.

XX N-PSDB; AAD27207.

XX Rat hepatoma H35 cell alpha1-2fucosyltransferase, useful for producing  
PT GM1-specific alpha1-2fucosyltransferase enzyme by recombinant  
PT techniques and for detecting oncogenic transformation of test tissues -

XX Claim 3; Fig 5; 41pp; English.

XX The invention relates to rat GM1-specific alpha1-2fucosyltransferase  
CC (alpha1-2Fuct) enzyme and its corresponding nucleic acid. This nucleic  
CC acid is specific for a carbohydrate moiety found in ganglioside GM1,  
CC a terminal galactose beta1-3N-acetylgalactosamine (Galbeta1-3GalNAc),  
CC saccharide. Alpha1-2Fuct DNA is useful for producing rat alpha1-2Fuct  
CC protein by recombinant techniques. Alpha1-2Fuct DNA is useful for the  
CC preparative synthesis of fucosyl containing glycolipids, glycoproteins,  
CC glycolipoproteins and oligosaccharide, and for preparing fucosyl-GM1.

CC Alpha-2Fuct DNA is useful for detecting oncogenic transformation which  
 CC involves assaying for changes in expression of alpha-2 Fuct. Since  
 CC alpha-2Fuct is activated in cell transformation, antisense sequences  
 CC derived from alpha-2Fuct DNA are useful for inhibiting, suppressing  
 CC or treating cancer. Alpha-2Fuct DNA is useful in gene therapy and  
 CC antisense therapy. The present sequence is rat hepatoma H35 cell  
 CC alpha-2Fuct.

XX Sequence 380 AA:

Query Match 100.0%; Score 1896; DB 23; Length 380;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-198;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQQRIVKLQPISEKELPMTOMSSGNTSEPMRDSQHGNGELRGMTINSIGRLGNOM 60  
 DB 28 LQQRIVKLQPISEKELPMTOMSSGNTSEPMRDSQHGNGELRGMTINSIGRLGNOM 87  
 QY 61 GEYATLFLARMNGRLAFIPASMHNALAPFRISLPVLSHDTAKKIPWNYHLNDWMEER 120  
 DB 88 GEYATLFLARMNGRLAFIPASMHNALAPFRISLPVLSHDTAKKIPWNYHLNDWMEER 147  
 QY 121 YRHIPGFHFVFTGYPCSWTFYHHLRPDLKEFTLHDHVRBAQAFGLRVNGSOPSTEV 180  
 DB 148 YRHIPGFHFVFTGYPCSWTFYHHLRPDLKEFTLHDHVRBAQAFGLRVNGSOPSTEV 207  
 QY 181 GYHVRGDIYHVPWPKGVADRGYIEKALDMFRARYSSPVVYVNSGMACRENTINAS 240  
 DB 208 GYHVRGDIYHVPWPKGVADRGYIEKALDMFRARYSSPVVYVNSGMACRENTINAS 267  
 QY 241 RGDVVFAGNIESSPAKDFALLTQCNTIMTIGTFGIMAYVLAGSDPIYANTLPDSPF 300  
 DB 268 RGDVVFAGNIESSPAKDFALLTQCNTIMTIGTFGIMAYVLAGSDPIYANTLPDSPF 327  
 QY 301 LKVFKEPAALFPEWVGIPADLSPLKALPACPRSHPHLNAKGVTCYVAGRAF 353  
 DB 328 LKVFKEPAALFPEWVGIPADLSPLKALPACPRSHPHLNAKGVTCYVAGRAF 380

RESULT 5  
 ID AAW37855 standard; Protein; 340 AA.

XX AAW37855;  
 XX  
 DT 28-AUG-1998 (first entry)  
 XX  
 DE Porcine secretor transferase (FUT2).  
 XX  
 KW Secretory; glycosyltransferase; FUT2; pig; epitope; antigen;  
 KM transgenic animal; xenotransplantation; organ transplant.  
 XX  
 OS Sus scrofa.

XX Key  
 FH Location/Qualifiers  
 FT Domain 1..4  
 FT Domain /note= "N-terminal cytoplasmic tail"  
 FT Domain 5..26  
 FT Domain /note= "transmembrane domain"  
 FT Domain 27..340  
 FT Domain /note= "C-terminal domain"  
 FT Modified-site 185..187  
 FT Modified-site /note= "Asn is N-glycosylated"  
 FT Modified-site 251..253  
 FT Modified-site /note= "Asn is N-glycosylated"  
 FT Modified-site 279..281  
 FT Modified-site /note= "Asn is N-glycosylated"  
 FT Modified-site 305..307  
 FT Modified-site /note= "Asn is N-glycosylated"

XX MO9807857-A1.  
 XX  
 XX  
 XX  
 PD 26-FEB-1998.

XX  
 PF 22-AUG-1997; 97NO-AU00540.  
 XX  
 PR 23-AUG-1996; 96AU-0001823.  
 XX  
 PA (AUST-) AUSTIN RES INST.  
 XX  
 PI McKenzie IFC, Sandrin MS;  
 XX  
 DR WPI; 1998-169148/15.  
 DR N-PSDB; AAV29003.

XX Nucleic acid encoding glycosyltransferase able to compete with  
 PT second such enzyme - particularly used to reduce expression of  
 PT unwanted carbohydrate epitope(s) on tissues intended for  
 PT transplantation  
 XX  
 PS Claim 6; Fig 1A-B; 40pp; English.

XX This polypeptide comprises porcine secretor glycosyltransferase  
 CC (SE or FUT2), a type II integral membrane protein has high affinity  
 CC for type I and type II substrates. Its amino acid sequence was  
 CC deduced from the nucleotide sequence of a genomic DNA clone (see  
 CC AAV29003) isolated from a pig liver library on the basis of homology  
 CC to human FUT2 cDNA. Pig FUT2 shows 83.2% identity with human FUT2.  
 CC Expression of a glycosyltransferase, such as porcine Se, in a  
 CC tissue results in reduced expression of unwanted carbohydrate  
 CC epitopes on the tissue, especially porcine heart, liver, kidney or  
 CC pancreas, rendering it more suitable for transplantation, i.e. less  
 CC immunogenic and of increased immunological acceptability. A  
 CC claimed method of producing a cell from a donor species that is  
 CC immunologically acceptable to a recipient species involves reducing  
 CC levels of carbohydrate on the donor cell that causes it to be  
 CC recognised as non-self by the recipient by expressing a nucleic  
 CC acid for a glycosyltransferase such as porcine Se in the cell.  
 XX

SO Sequence 340 AA:

Query Match 75.4%; Score 1430; DB 19; Length 340;  
 Best Local Similarity 81.9%; Pred. No. 1.2e-147;  
 Matches 267; Conservative 26; Mismatches 19; Indels 14; Gaps 3;

QY 1 LQQRIVKLQPISEKELPMTOMSSGNTSEPMRDSQHGNGELRGMTINSIGRLGNOM 60  
 DB 28 LQQRIVKLQPISEKELPMTOMSSGNTSEPMRDSQHGNGELRGMTINSIGRLGNOM 73  
 QY 61 GEYATLFLARMNGRLAFIPASMHNALAPFRISLPVLSHDTAKKIPWNYHLNDWMEER 120  
 DB 74 GEYATLFLARMNGRLAFIPASMHNALAPFRISLPVLSHDTAKKIPWNYHLNDWMEER 133  
 QY 121 YRHIPGFHFVFTGYPCSWTFYHHLRPDLKEFTLHDHVRBAQAFGLRVNGSOPSTEV 180  
 DB 134 YRHIPGFHFVFTGYPCSWTFYHHLRPDLKEFTLHDHVRBAQAFGLRVNGSOPSTEV 193  
 QY 181 GYHVRGDIYHVPWPKGVADRGYIEKALDMFRARYSSPVVYVNSGMACRENTINAS 240  
 DB 194 GYHVRGDIYHVPWPKGVADRGYIEKALDMFRARYSSPVVYVNSGMACRENTINAS 253  
 QY 241 RGDVVFAGNIESSPAKDFALLTQCNTIMTIGTFGIMAYVLAGSDPIYANTLPDSPF 300  
 DB 254 RGDVVFAGNIESSPAKDFALLTQCNTIMTIGTFGIMAYVLAGSDPIYANTLPDSPF 313  
 QY 301 LKVFKEPAALFPEWVGIPADLSPLK 326  
 DB 314 LKVFKEPAALFPEWVGIPADLSPLK 339

RESULT 6  
 ID AAW53101 standard; Protein; 340 AA.

XX AAW53101;  
 XX  
 XX

```

DT 08-JUL-1998 (first entry)
XX Pig secretor.
XX Pig; secretor; chimeric; glycosyltransferase; gene therapy;
XX transplantation.
XX Sus scrofa.
XX WO9805768-A1.
XX 12-FEB-1998.
XX 01-AUG-1997; 97WO-AU00482.
XX 21-AUG-1996; 96US-0024279.
XX 02-AUG-1996; 96AU-0001402.
XX (AUST-) AUSTIN RES INST.
XX McKenzie IFC, Sandrin MS;
XX WPI; 1998-159170/14.
XX N-PSDB; AAV21639.
XX Nucleic acids encoding chimeric glycosyltransferases - used for
XX altering carbohydrate levels on the surface of cells, useful in gene
XX therapy and transplantation
XX Example 2; Fig 6; 51pp; English.
XX The present sequence represents pig secretor used in an example of the
XX present invention. The present invention describes nucleic acids (NA)
XX encoding a chimeric glycosyltransferase. The NA comprises a catalytic
XX domain of a first glycosyltransferase (GT) and a localisation signal of
XX a second GT, whereby when the NA is expressed in a cell and where the
XX chimeric enzyme is located in an area of the cell where it is able to
XX compete for substrate with a second GT, resulting in reduced levels of
XX a product from the second GT. The NA can be used to produce cells and
XX organs with desired glycosylation patterns. Products and methods of the
XX present invention can be used to reduce the levels of undesirable
XX epitopes in cells, tissues or organs which may be used in
XX transplantation or gene therapy.
XX Sequence 340 AA;
XX Query Match 75.4%; Score 1430; DB 19; Length 340;
XX Best Local Similarity 81.9%; Pred. No. 1.2e-147;
XX Matches 267; Conservative 26; Mismatches 19; Indels 14; Gaps 3;
QY 1 LQORIVKLQPLSEKELPMTTOMSSGNTESPERRDSECHNGCELGRMFTINSIGRLGNQM 60
Db 28 LQORIVKLQPLSEKELPMTTOMSSGNTESPERRDSECHNGCELGRMFTINSIGRLGNQM 73
QY 61 GEYATLFLALARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
Db 74 GEYATLFLALARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPWQNYHLNDWMEER 133
QY 121 YRHIPGFHVFTGYPCSWTFFYHLLRPEILKEFTLHDHVREAAQFLRGLRVNGSQPSTFV 180
Db 134 YRHIPGFHVFTGYPCSWTFFYHLLRPEILKEFTLHDHVREAAQFLRGLRVNGSQPSTFV 193
QY 181 GVHVRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSPFVVTNSGMACWCRINAS 240
Db 194 GVHVRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSPFVVTNSGMACWCRINAS 253
QY 241 RGVVVFAGNGTEGSPAKDFALLTQCNHTIMTIGTFGWAAYLAGDGIYLYANTYLPDSPF 300
Db 254 RGVVVFAGNGTEGSPAKDFALLTQCNHTIMTIGTFGWAAYLAGDGIYLYANTYLPDSPF 313
QY 301 LKVEKPEAFLPEWVGIPADLSPLK 326
Db 314 LKVEKPEAFLPEWVGIPADLSPLK 339

```

```

RESULT 7
AAW69332
ID AAW69332 standard; Protein: 343 AA.
XX AC AAW69332;
XX 20-NOV-1998 (first entry)
XX Human Sec2 protein sequence.
XX Sec2; alpha(1,2) fucosyltransferase; H blood group; secretor genotyping;
XX KW GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase; human;
XX KW FUT2; nonsecretor genotyping.
XX OS Homo sapiens.
XX US5807732-A.
XX 15-SEP-1998.
XX 28-FEB-1995; 95US-0395800.
XX 28-FEB-1995; 95US-0395800.
XX (GIOR/) GIORGI D.
XX (KELL/) KELLY R. J.
XX (LENN/) LENNON G.
XX (LOWE/) LOWE J. B.
XX (ROUQ/) ROUQUIER S.
XX Giorgi D, Kelly RJ, Lennon G, Lowe JB, Rouquier S;
XX WPI; 1998-520127/44.
XX N-PSDB; AAV58323.
XX DNA encoding fucosyltransferase enzyme - useful for producing
XX recombinant enzyme and genotyping person as secretor or nonsecretor
XX Disclosure; Column 45-50; 55pp; English.
XX This sequence is the human Sec2 protein of the invention. The DNA
XX encodes a alpha(1,2) fucosyltransferase and is the Secretor
XX alpha(1,2) fucosyltransferase locus, that cross hybridises with the
XX H blood group alpha(1,2) fucosyltransferase gene. The DNA is useful for
XX producing a recombinant human GDP-L-fucose:beta-D-galactoside
XX 2-alpha-L-fucosyltransferase (FUT2) which can be used for genotyping an
XX individual as a secretor or nonsecretor as it is known that nonsecretors
XX homozygous for a mutant allele of the FUT2 gene that has a stop codon in
XX the position corresponding to amino acid 143.
XX Sequence 343 AA;
XX Query Match 73.1%; Score 1386.5; DB 19; Length 343;
XX Best Local Similarity 79.1%; Pred. No. 7.3e-143;
XX Matches 256; Conservative 26; Mismatches 31; Indels 11; Gaps 2;
QY 1 LQORIVKLQPLSEKELPMTTOMSSGNTESPERRDSECHNGCELGRMFTINSIGRLGNQM 60
Db 28 VQQLAKTIQAM--WELPV-----QIPVLASTSKALGPSQLRGMTWTINAIKRLGNQM 76
QY 61 GEYATLFLALARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
Db 77 GEYATLFLALARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPWQNYHLNDWMEER 136
QY 121 YRHIPGFHVFTGYPCSWTFFYHLLRPEILKEFTLHDHVREAAQFLRGLRVNGSQPSTFV 180
Db 137 YRHIPGFHVFTGYPCSWTFFYHLLRPEILKEFTLHDHVREAAQFLRGLRVNGSQPSTFV 196
QY 181 GVHVRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSPFVVTNSGMACWCRINAS 240
Db 197 GVHVRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSPFVVTNSGMACWCRINAS 256

```

QY 241 RGDVVFAGNIEGSPAKDFALLTQCNHITMTIGTIGIWAAYLAGDITILANYTLPDSPF 300  
 DB 257 RGDVVFAGNIEGSPAKDFALLTQCNHITMTIGTIGIWAAYLAGDITILANYTLPDSPF 316  
 QY 301 LKVFPEAAFLPEWVGIPADLSPLK 326  
 DB 317 LKVFPEAAFLPEWVGIPADLSPLK 342

## RESULT 8

AAB36106  
 ID AAB36106 standard; Protein; 344 AA.  
 AC AAB36106;

DT 19-FEB-2001 (first entry)

XX Human Sec2 catalytic domain.

XX Human; Sec2; alpha1-2fucosyltransferase; cytotstatic;  
 KW neuroprotective; nootropic; gene therapy; Fucalpalhal-2galbatal-3galNAC;  
 KW immunotherapy; immunosuppression; cancer; neurological disease;  
 KW small cell lung carcinoma.

OS Homo sapiens.

XX WO200064464-A1.

PN 02-NOV-2000.

XX 23-APR-1999; 99WO-US07384.

XX 23-APR-1999; 99WO-US07384.

PA (PACI-) PACIFIC NORTHWEST CANCER FOUND.

XX Holmes EH, Sherwood AL;

XX WPI; 2000-687262/67.

XX New rat ganglioside GM1-specific alpha1-2fucosyltransferase, useful for  
 PT preparation of fucosyl GM1 which is useful as a nutritional composition  
 PT or immunotherapeutic for cancer and neurological diseases.

XX Example; Fig 3B; 91pp; English.

XX The present sequence is given in a specification relating to a rat  
 CC ganglioside GM1-specific alpha1-2fucosyltransferase protein. The protein  
 CC or its cellular fraction is useful for synthesis of a molecule comprising  
 CC Fucalpalhal-2galbatal-3galNAC, a glycolipid, glycoprotein, glycolipoprotein  
 CC or a free oligosaccharide comprising Fucalpalhal-2galbatal-3galNAC.  
 CC The method involves contacting alpha1-2fucosyltransferase with GDP-fucose  
 CC and a molecule or glycolipid, glycoprotein, glycolipoprotein or  
 CC oligosaccharide having a terminal Galbatal-3galNAC group. It is also  
 CC useful for synthesis of fucosyl-GM1 by contacting the protein with  
 CC GDP-fucose and ganglioside GM1. The obtained glycoproteins,  
 CC glycolipoproteins, glycolipids and oligosaccharides are useful as  
 CC nutritional compositions and fucosyl-GM1 is useful for inducing an  
 CC immunotherapeutic or immunosuppressive action against cancer,  
 CC neurological disease or small cell lung carcinoma.

XX Sequence 344 AA;

Query Match 72.6%; Score 1376; DB 21; Length 344;  
 Best Local Similarity 78.9%; Pred. No. 1e-141;  
 Matches 258; Conservative 26; Mismatches 31; Indels 12; Gaps 3;

QY 1 LQRIYKLOPLSEKELPMTOMSSGNTESEPMRDSQHGNGELRGMTINSTIGRLGNOM 60

DB 28 VQORLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGMTINAIKRLGNOM 76

QY 61 GEYATLALARMNGRLAFTAPSMHNAALAFIRISLFLVLSHSDTAKKIPWNYHLNDWKEE 120

DB 77 GEYATLALAKWNGREAPFPAQMHSTLAPFIRIILPVLSHATASRIPQWNYHLNDWKEE 136  
 QY 121 YRHI-PGHEVVRITGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGRVNGSPSTF 179  
 DB 137 YRHIIPGGEYVRITGYPCSWTFYHHLRQELQFTLHDHVREAAQAFRLGRVNGSPSTF 196  
 QY 180 VGVHVRGDIYVHMPNVKGVVADRGYLEKALDMFRARYSSPVVTSNGMAWCRENINA 239  
 DB 197 VGVHVRGDIYVHMPKVKGVVADRRYLOQALDWFARYSSLIFFVTSNGMAWCRENIDT 256  
 QY 240 SRGDVVFAGNIEGSPAKDFALLTQCNHITMTIGTIGIWAAYLAGDITILANYTLPDSP 299  
 DB 257 SRGDVVFAGNIEGSPAKDFALLTQCNHITMTIGTIGIWAAYLAGDITILANYTLPDSP 316  
 QY 300 ELKVFPEAAFLPEWVGIPADLSPLK 326  
 DB 317 ELKVFPEAAFLPEWVGIPADLSPLK 343

## RESULT 9

AAR13751

ID AAR13751 standard; Protein; 365 AA.

XX AAR13751;

DT 07-NOV-1991 (first entry)

DE GDP-Fuc:beta-D-galactoside alpha(1,2)-fucosyltransferase.

XX Glycosyltransferase.

OS Homo sapiens.

PN WO9112340-A.

XX 22-AUG-1991.

XX 14-FEB-1991; 91WO-US00899.

XX 12-DEC-1990; 90US-0627621.

XX 14-FEB-1990; 90US-0479858.

XX 14-FEB-1990; 90US-0480133.

PA (UNMI ) UNIV OF MICHIGAN.

XX Lowe JB;

XX WPI; 1991-267151/36.

DR P-PSDB; AAR13751.

XX Isolation of gene conveying post-translational characteristic -  
 PT e.g. the presence of soluble or membrane bound oligo or  
 PT polysaccharide or glycosyltransferase.

XX Disclosure; Fig 3; 155pp; English.

XX The amino acid sequence codes for a protein capable of functioning  
 CC as a GDP-Fuc:(beta-D-Gal alpha(1,2))-fucosyltransferase. The  
 CC functional protein is represented by amino acids 33 to 365. The  
 CC enzyme produced by the DNA sequence can be used in enzymatic  
 CC fucosylation of chain-terminating galactose residues on lactose-  
 CC amine or neolacto type beta-D-galactoside to alpha-2-L-fucose  
 CC residues. See also AAR13749-R13752.

XX Sequence 365 AA;

Query Match 55.5%; Score 1051.5; DB 12; Length 365;  
 Best Local Similarity 63.3%; Pred. No. 4e-106;  
 Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

QY 25 GNTSEPMRDSQHGNGELRGMTINSTIGRLGNOMGEYATLALARMNGRLAFTAPSHH 84

```

Db 59 GTAMGNASSCPQH-PASLSGTWTVPNGRFGNQMGQVATLLALAQLNGRRRAFIIPAMH 117
QY 85 NALAPIERISLPVLSHSDAKKTPWQNYHLNDWMEERYRHIPGFVFTGYPGCSWTFYHHL 144
Db 118 AALAPVFRITLVLAPEDVSRTPWRELQHDWMESEYADLRDPFLKLSGFPSCWTFYHHL 177
QY 145 RPEILKEFTLHDHVRREAAQAFRLGRV--NGSQPSTFVGHVHVRGDDYVHMVPMVKGVA 202
Db 178 REQIRREFTLHDHVRREAAQAFRLGRV--NGSQPSTFVGHVHVRGDDYVHMVPMVKGVA 237
QY 203 DRGYLEKALDMFRARYSSPVFVVTNSGMAWCRENINASRGDVVVFAGNIEGSPAKDFALL 262
Db 238 DSAYLRQAMDWRARHEAPVFFVVTNSGMECKENITSGDVTFAAGDQEAATPWKDFALL 297
QY 263 TCQNTHTIMTIGTGGWAAYLAGDGTIYLANFTLPDSFELKIFKPEAAFLPEWVGIPADLS 322
Db 298 TCQNTHTIMTIGTGGWAAYLAGDGTIYLANFTLPDSFELKIFKPEAAFLPEWVGIPADLS 357
QY 323 PLLKALTP 330
Db 358 PLWTLAKP 365

RESULT 10
ID AAR80154 standard; Protein; 365 AA.
XX
AC AAR80154;
XX
XX 25-JUN-1996 (first entry)
XX
DE GDP-L-fucose-beta-D-galactosidase alpha-1,2-fucosyltransferase.
XX
KW GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyl-transferase; Fuc-T;
KW synthesis; 2'-fucosyllactose; oligosaccharide; human milk;
KW non-human transgenic mammal; secondary gene product.
XX
OS Homo sapiens.
XX
PN W09524494-A1.
XX
PD 14-SEP-1995.
XX
PF 24-JAN-1995; 95WO-US00926.
XX
PR 09-MAR-1994; 94US-0209122.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Cummings RD, Kopchik JJ, Moreman KW, Mukerji P;
PI Pierce JM, Prieto PA, Smith DF;
XX
DR WPI: 1995-336739/43.
XX
N-PSDB; AAQ98461.
XX
PT Prodn. of human:ised milk by non-human transgenic mammal - by
PT inserting heterologous gene encoding human catalytic entity which
PT produces oligo:saccharide(s) and glyco:conjugate(s) into mammal's
PT genome
XX
PS Example 1; Page 62-64; 83pp; English.
XX
XX GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyl-transferase (Fuc-T)
CC is encoded by cDNA (AAQ98461) isolated from a human epidermal carcinoma
CC cell line. The enzyme is responsible for the synthesis of
CC 2'-fucosyllactose, one of the oligosaccharides in human milk. Other
CC products of Fuc-T include glycoproteins contg. beta-linked terminal
CC galactose residues which can be fucosylated by Fuc-T. This DNA can
CC be used to produce non-human transgenic mammals able to produce
CC secondary gene products, e.g. oligosaccharides, in their milk. The
CC transgenic mammals milk biochemically resembles human milk. This
CC humanised milk can be used in the prepn. of an enteral nutritional
CC product useful in the nutritive maintenance of an animal.

```

```

XX Sequence 365 AA:
QY 25 GNTGSEPMERDSEQHGNGELRGMTINSTIGRLGNQMGQVATLLALAQLNGRRRAFIIPAMH 84
Db 59 GTAMGNASSCPQH-PASLSGTWTVPNGRFGNQMGQVATLLALAQLNGRRRAFIIPAMH 117
QY 85 NALAPIERISLPVLSHSDAKKTPWQNYHLNDWMEERYRHIPGFVFTGYPGCSWTFYHHL 144
Db 118 AALAPVFRITLVLAPEDVSRTPWRELQHDWMESEYADLRDPFLKLSGFPSCWTFYHHL 177
QY 145 RPEILKEFTLHDHVRREAAQAFRLGRV--NGSQPSTFVGHVHVRGDDYVHMVPMVKGVA 202
Db 178 REQIRREFTLHDHVRREAAQAFRLGRV--NGSQPSTFVGHVHVRGDDYVHMVPMVKGVA 237
QY 203 DRGYLEKALDMFRARYSSPVFVVTNSGMAWCRENINASRGDVVVFAGNIEGSPAKDFALL 262
Db 238 DSAYLRQAMDWRARHEAPVFFVVTNSGMECKENITSGDVTFAAGDQEAATPWKDFALL 297
QY 263 TCQNTHTIMTIGTGGWAAYLAGDGTIYLANFTLPDSFELKIFKPEAAFLPEWVGIPADLS 322
Db 298 TCQNTHTIMTIGTGGWAAYLAGDGTIYLANFTLPDSFELKIFKPEAAFLPEWVGIPADLS 357
QY 323 PLLKALTP 330
Db 358 PLWTLAKP 365

RESULT 11
AAR70422
ID AAR70422 standard; Protein; 365 AA.
XX
AC AAR70422;
XX
XX 27-FEB-1996 (first entry)
XX
DE 2-Alpha-fucosyltransferase.
XX
KW GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyltransferase;
KW glycosyltransferase; glycosylation; oligosaccharide; glycoprotein;
KW glycolipid; transgenic animal; cattle; milk.
XX
OS Homo sapiens.
XX
PN W09524495-A1.
XX
PD 14-SEP-1995.
XX
PF 24-JAN-1995; 95WO-US00967.
XX
PR 09-MAR-1994; 94US-0208889.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Cummings RD, Kopchik JJ, Moreman KW, Mukerji P;
PI Pierce JM, Prieto PA, Smith DF;
XX
DR WPI: 1995-328284/42.
XX
N-PSDB; AAT01083.
XX
PT New transgenic non-human mammal milk prods - contg. heterologous
XX components produced as secondary gene prods. of an heterologous gene
XX Example 1; Page 62-64; 83pp; English.
XX
XX 2-Alpha-fucosyltransferase (AAR70422) is encoded by a cDNA clone
CC (AAT01083) isolated from a human epidermal carcinoma A431 library.
CC The enzyme can be expressed in the milk of a transgenic mammal,
CC esp. cow. This allows large-scale prodn. of oligosaccharides and

```

CC glycosylated proteins and lipids in the milk.

XX Sequence 365 AA;

Query Match 55.5%; Score 1051.5; DB 16; Length 365;  
Best Local Similarity 63.3%; Pred. No. 4e-106;  
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

QY 25 GNTSEPMRDRSEHGNGELRCMFTINSIGRLGNOMGEYATLALARMNGRLAFIPASMH 84  
DB 59 GTAMGPNASSCPQH-PASLSGTWTVYVNGRFGNMGQYAILLALAQNGRAFIIPAMH 117  
QY 85 NALAPTRISLPLVLSHTAKKIPQWNYHLNDWMEERYRHIPGHVFRFTGYPCSWTFYHHL 144  
DB 118 AALAPVFRITLVLAPVDSRTPEWRELQLDHNMSEYADLRDPFLKSGFPCSWTFYHHL 177  
QY 145 REILKEFTLHDHREEAQAFGLRV--NGSQPSTFVGVHVRGDDYVHVMNVKGVVA 202  
DB 178 REQIRREFTLHDHREEAQSVLGQLRGTRGDRPTFVGVHVRGDDYVHVMNVKGVVG 237  
QY 203 DRGYLEKALDMFRARYSSPVVYVTSNGMAWCRENINASRGDVVFNAGNTEGSPAKDFALL 262  
DB 238 DSAYLQAMDFRAREAPVYVTSNGMEWCKENIDTSQGDVTFAGDGOEATPWKDFALL 297  
QY 263 TCNHTIMTIGTFGIWAAYLAGDITVILANYTLPSFPLKVKPEAAFLPEWVGIPADLS 322  
DB 298 TCNHTIMTIGTFGIWAAYLAGDITVILANYTLPSFPLKVKPEAAFLPEWVGIPADLS 357  
QY 323 PLLKALTP 330  
DB 358 PLWTLAKP 365

RESULT 12

AAR70421  
ID AAR70421 standard; Protein: 365 AA.

XX AAR70421;

XX 27-FEB-1996 (first entry)

XX 2-Alpha-fucosyltransferase.

XX GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyltransferase;  
KW glycosyltransferase; oligosaccharide; glycoprotein; glycolipid;  
KW transgenic animal; cattle; glycosylation; milk.

XX Homo sapiens.

XX WO9524488-A1.

XX 14-SEP-1995.

XX 24-JAN-1995; 95WO-US01147.

XX 09-MAR-1994; 94US-0209132.

XX (ABBO ) ABBOTT LAB.

XX Cummings RD, Kopechik JJ, Moremen KW, Mukerji P;

XX Pierce JW, Prieto PA, Smith DF;

XX WPI; 1995-328279/42.

XX N-PSDB; AAT01082.

XX Transgenic animal expressing heterologous catalyst - used in  
PT metabolite prodn esp. glycosyltransferase for prodn. of  
PT oligosaccharide(s) and glyco-conjugate(s) in the milk, useful in  
PT foods, pharmaceuticals, etc.

XX Example 1; Page 62-64; 84pp; English.

XX Alpha-1,2-fucosyltransferase (AAR70421) is encoded by a cDNA clone

CC (AAT01082) isolated from a human epidermal carcinoma cell line A431  
CC cDNA library. The enzyme can be expressed in the milk of a  
CC transgenic mammal, esp. cow. This allows large-scale prodn. of  
CC oligosaccharides or glycosylated proteins and lipids in the milk.

XX Sequence 365 AA;

Query Match 55.5%; Score 1051.5; DB 16; Length 365;  
Best Local Similarity 63.3%; Pred. No. 4e-106;  
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

QY 25 GNTSEPMRDRSEHGNGELRCMFTINSIGRLGNOMGEYATLALARMNGRLAFIPASMH 84  
DB 59 GTAMGPNASSCPQH-PASLSGTWTVYVNGRFGNMGQYATLALAQNGRAFIIPAMH 117  
QY 85 NALAPTRISLPLVLSHTAKKIPQWNYHLNDWMEERYRHIPGHVFRFTGYPCSWTFYHHL 144  
DB 118 AALAPVFRITLVLAPVDSRTPEWRELQLDHNMSEYADLRDPFLKSGFPCSWTFYHHL 177  
QY 145 REILKEFTLHDHREEAQAFGLRV--NGSQPSTFVGVHVRGDDYVHVMNVKGVVA 202  
DB 178 REQIRREFTLHDHREEAQSVLGQLRGTRGDRPTFVGVHVRGDDYVHVMNVKGVVG 237  
QY 203 DRGYLEKALDMFRARYSSPVVYVTSNGMAWCRENINASRGDVVFNAGNTEGSPAKDFALL 262  
DB 238 DSAYLQAMDFRAREAPVYVTSNGMEWCKENIDTSQGDVTFAGDGOEATPWKDFALL 297  
QY 263 TCNHTIMTIGTFGIWAAYLAGDITVILANYTLPSFPLKVKPEAAFLPEWVGIPADLS 322  
DB 298 TCNHTIMTIGTFGIWAAYLAGDITVILANYTLPSFPLKVKPEAAFLPEWVGIPADLS 357  
QY 323 PLLKALTP 330  
DB 358 PLWTLAKP 365

RESULT 13

AAR90572

ID AAR90572 standard; Protein: 365 AA.

XX AAR90572;

XX 08-APR-1996 (first entry)

XX Human H-transferase.

XX H-transferase; xenograft hyperacute rejection; transplantation;  
KW glycosyltransferase; galactose alpha(1,3) galactose.

XX Homo sapiens.

XX WO9534202-A1.

XX 21-DEC-1995.

XX 14-JUN-1995; 95WO-US07554.

XX 21-JUL-1994; 94US-0278282.

XX 15-JUN-1994; 94US-0260201.

XX (ALEX-) ALEXION PHARM INC.

XX (AUST-) AUSTIN RES INST.

XX Fodor WL, McKenzie IFC, Rother RP, Sandrin MS; Squinto SP;

XX WPI; 1996-049326/05.

XX N-PSDB; AAT12238.

XX Redn. of rejection of xenogeneic cells following transplantation  
PT by introducing a vector expressing fucosyl:transferase into the  
PT cells

XX Example 1; Page 45-47; 69pp; English.

```

XX CC The human H-transferase (AA390572) product of a cDNA clone (AAT12238);
CC prep. from human epidermoid carcinoma cells (ATCC CRL 155 A-431)
CC can be expressed in xenogeneic organs, tissues and cells using
CC a vector such as pAPEX-1 (AAT12239). This results in decreased
CC expression of the non-human antigen galactose alpha(1,3) galactose
CC on the surface of the organs etc. so that hyperacute rejection is
CC reduced upon transplantation to humans.
XX CC
XX Sequence 365 AA;
XX
Query Match 55.5%; Score 1051.5; DB 17; Length 365;
Best Local Similarity 63.3%; Pred. No. 4e-106;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 25 GNTESPMRDRSEQHNGELRGMTINSIGRLGNOMGEYATLFLARNMGRLAFTIPASMH 84
DB 59 GTAMGNPNASSSCPQH-PASLSGTTWVYPNGRFGNMGQYATLLAQLNGRRAPILPAMH 117
QY 85 NALAPITRISLPVLSHTAKKIPQNYHLNDWMEERYRHPGHFVRFTGYPCSWTFYHHL 144
DB 118 AALAPVRIITLPVLAPEVDSKTPWRELQLDHMKSEYADLRDPFLKLSGPPCSWTFEHL 177
QY 145 RPEILKFTLHDHVREREAQAFRLGRV--NGSQPSTFVGVHVRGDDYVHMPNWKGVYA 202
DB 178 REQIRREFTLHDHLREEAQSVLGQLRLGRTGDRPRTFVGVHVRGDDYLVQVMPQWKGVG 237
QY 203 DRGYLEKALDMFRARYSSPVFVTSNGMWCRENINASRGDVFAGNGIEGSPAKDFALL 262
DB 238 DSAYLRQAMDWFRARHEAPFVVTSGMEWCKENIDTSQGDVTFAGDGOEATPWKDFALL 297
QY 263 TCNHTTMTIGTFGIWAAYLAGGDTIYLANVTLDPSPFLKVKPEAAFLPEWVGIPADLS 322
DB 298 TCNHTTMTIGTFGFWAAYLAGGDTVYLANVTLDPSEFLKFKPEAAFLPEWVGINADLS 357
QY 323 PLLKALTP 330
DB 358 PLWTIAKP 365
RESULT 14
AAW23805
ID AAW23805 standard; Protein; 365 AA.
XX AC
XX AAW23805;
XX
DT 15-SEP-1997 (first entry)
XX
XX Human alpha 1,2 fucosyltransferase (EC 2.4.1.69).
XX
XX Human GDP-L-fucose:beta-D-galactoside-2-alpha-L-fucosyltransferase;
KW alpha 1,2 FT; transgene; transgenic mouse; animal model;
KW intestinal adhesion; Helicobacter pylori infection; stomach;
KW small intestine; gut; epithelial cell; surface receptor;
KW carbohydrate antigen; gastritis; peptic ulcer; neoplasia;
KW gastric adenocarcinoma; Lewis antigen; fucosylation.
XX
XX Homo sapiens.
XX
XX US5625124-A.
XX
XX 29-APR-1997.
XX
XX 11-JUL-1994; 94US-0273411.
XX
XX 11-JUL-1994; 94US-0273411.
XX
XX (UNIW ) UNIV WASHINGTON.
XX
XX Falk P, Gordon JI;
XX
XX WPI; 1997-258275/23.
XX
XX N-PSDB; AAT76768.

```

```

XX PT Animal model for Helicobacter pylori infection - comprising
XX transgenic mouse expressing human enzyme promoting intestinal
XX adhesion
XX
XX Example 1; Columns 13-16; 24pp; English.
XX
XX A claimed transgenic mouse expresses, in its intestinal epithelial
XX cells, the enzyme human GDP-L-fucose:beta-D-galactoside 2-alpha-L-
XX fucosyltransferase (also called alpha 1,2 FT) or human GDP-L-fucose:
XX beta-D-N-acetylglucosamide 3/4-alpha-L-fucosyltransferase (also
XX called alpha 1,3/4 FT). The enzyme is expressed under the
XX control of a gut epithelial cell-specific promoter and Helicobacter
XX pylori adheres to the transgenic cells. The transgenic mouse and
XX intestinal epithelial cells from it are useful as models for screening
XX compounds for the ability to inhibit adhesion of H. pylori to gut
XX epithelial cells. The present sequence represents human alpha 1,2 FT
XX and was published by Larsen et al., Proc.Natl.Acad.Sci.USA, 87:
XX 6674-6678 (1990).
XX
XX Sequence 365 AA;
XX
Query Match 55.5%; Score 1051.5; DB 18; Length 365;
Best Local Similarity 63.3%; Pred. No. 4e-106;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 25 GNTESPMRDRSEQHNGELRGMTINSIGRLGNOMGEYATLFLARNMGRLAFTIPASMH 84
DB 59 GTAMGNPNASSSCPQH-PASLSGTTWVYPNGRFGNMGQYATLLAQLNGRRAPILPAMH 117
QY 85 NALAPITRISLPVLSHTAKKIPQNYHLNDWMEERYRHPGHFVRFTGYPCSWTFYHHL 144
DB 118 AALAPVRIITLPVLAPEVDSKTPWRELQLDHMKSEYADLRDPFLKLSGPPCSWTFEHL 177
QY 145 RPEILKFTLHDHVREREAQAFRLGRV--NGSQPSTFVGVHVRGDDYVHMPNWKGVYA 202
DB 178 REQIRREFTLHDHLREEAQSVLGQLRLGRTGDRPRTFVGVHVRGDDYLVQVMPQWKGVG 237
QY 203 DRGYLEKALDMFRARYSSPVFVTSNGMWCRENINASRGDVFAGNGIEGSPAKDFALL 262
DB 238 DSAYLRQAMDWFRARHEAPFVVTSGMEWCKENIDTSQGDVTFAGDGOEATPWKDFALL 297
QY 263 TCNHTTMTIGTFGIWAAYLAGGDTIYLANVTLDPSPFLKVKPEAAFLPEWVGIPADLS 322
DB 298 TCNHTTMTIGTFGFWAAYLAGGDTVYLANVTLDPSEFLKFKPEAAFLPEWVGINADLS 357
QY 323 PLLKALTP 330
DB 358 PLWTIAKP 365
RESULT 15
AAW13640
ID AAW13640 standard; Protein; 365 AA.
XX AC
XX AAW13640;
XX
XX 19-JUN-1997 (first entry)
XX
XX Human alpha(1,2)-fucosyltransferase.
XX
XX Alpha(1,2)-fucosyltransferase; glycosylation; oligosaccharide;
KW blood group H.
XX
XX Homo sapiens.
XX
XX WO9709421-A1.
XX
XX 13-MAR-1997.
XX
XX 06-SEP-1996; 96WO-US13816.
XX
XX 08-SEP-1995; 95US-0525058.
XX
XX PR

```

```
XX (UNMI ) UNIV MICHIGAN.
XX PA
XX PI Legault DJ, Lowe JB;
XX DR WPI: 1997-192897/17.
XX N-PSDB; AAT61677.
XX PT New recombinant fucosyltransferase proteins - useful for modifying
XX cell surface oligosaccharide structures
XX PS
XX Example 1; Page 280-281; 329pp; English.
XX Human GDP-Fuc:beta-D-galactoside alpha(1,2)-fucosyltransferase
XX (AAW13640) is an enzyme involved in the expression of type I and II
XX blood group H structures. Its amino acid sequence was deduced
XX from an isolated DNA (AAT61677) derived from human A431 cells.
XX Expression of the fucosyltransferase animal cell lines provides
XX specific capabilities with respect to post-translational
XX modification of the oligosaccharides of expressed proteins or
XX lipids. Specific applns. of the enzyme include enzymatic
XX fucosylation of chain-terminating galactose residues on
XX lactosamine or neolacto type beta-D-galactoside to alpha-2-
XX fucose residues. The enzyme can also be used to raise antibodies
XX as diagnostic reagents and to screen cpds. for fucosyltransferase
XX inhibitor activity.
XX SQ Sequence 365 AA;
XX
Query Match 55.5%; Score 1051.5; DB 18; Length 365;
Best Local Similarity 63.3%; Pred. No. 4e-106;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 25 GNTSEPMRRSEQHNGELRGMTINSIGRLGNOMGEYAILFALARMNGRLAFIPASMH 84
DB 59 GTAMGNASSSCPQH-PASLSGTWTVPYNGRFGNMGQYATLLAQLNGRAFIIPAMH 117
QY 85 NALAPIRISLPLHSDTAKKIPWNYHLNDWMEERYRHIPGCHVFRFTGYPCSWTFYHHL 144
DB 118 AALAPVFRITLVLAPVDSTRPWEQLQHDWMSSEYADLRDPFLKLSGFCPSWTFYHHL 177
QY 145 RPEILKEFTLHDHVRREAQAFRLGRV--NGSQPSTFVGHVHVRGDIYHVMPNWKGVYA 202
DB 178 REQIRREFTLHDHREEAQSVLGQLRLGRTGDRPRTFVGHVHVRGDIYLVQMPQRWKGVYG 237
QY 203 DRGYLEKALDMFRARYSSPVVYVTSNGMAWCARENINASEGDVVFAGNGIEGSPAKDFALI 262
DB 238 DSAYLRQAMDFRAREHAPVYVTSNGMEWCKENICTSGQDVTFAQDGOEATPWKDFALL 297
QY 263 TCQNTHTIMTIGTGIWAAYLAGDITIYLANITLPDPSFLKFKPEAAFLPENWVGIPADLS 322
DB 298 TCQNTHTIMTIGTGGWAAYLAGDVIYLANITLPDSEFLKFKPEAAFLPENWGINADLS 357
QY 323 PLLKALTP 330
DB 358 PLWTLAKP 365
```

Search completed: May 27, 2003, 15:06:16  
Job time : 45.7707 secs



Wed May 28 09:21:51 2003

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:03:22 ; Search time 14.5136 Seconds  
(without alignments)  
715.625 Million cell updates/sec

Title: US-10-040-863-10  
Perfect score: 1896  
Sequence: 1 LQORIVKLOPLSEKELPMT.....RSHPHLXAKGVTCYVAGRAF 353

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA: \*  
1: /cgn2.6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2.6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2.6/ptodata/1/1aa/6A\_COMB.pep.\*  
4: /cgn2.6/ptodata/1/1aa/6B\_COMB.pep.\*  
5: /cgn2.6/ptodata/1/1aa/6C\_COMB.pep.\*  
6: /cgn2.6/ptodata/1/1aa/6D\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1896	100.0	333	US-09-298-886-10	Sequence 10, Appl
2	1896	100.0	380	US-09-298-886-8	Sequence 8, Appl
3	1430	75.4	340	US-09-254-077A-6	Sequence 6, Appl
4	1430	75.4	340	US-09-254-077A-7	Sequence 7, Appl
5	1386.5	73.1	343	US-08-395-800A-8	Sequence 8, Appl
6	1386.5	73.1	343	US-09-254-077A-8	Sequence 8, Appl
7	1376	72.6	344	US-09-298-886-11	Sequence 11, Appl
8	1375.5	72.5	347	US-09-254-077A-9	Sequence 9, Appl
9	1080.5	57.0	373	US-09-254-077A-12	Sequence 12, Appl
10	1051.5	55.5	333	PCT-US91-00899-11	Sequence 11, Appl
11	1051.5	55.5	365	US-07-914-281-6	Sequence 6, Appl
12	1031.5	55.5	365	US-08-393-246-6	Sequence 6, Appl
13	1051.5	55.5	365	US-08-273-411-1	Sequence 1, Appl
14	1051.5	55.5	365	US-08-525-058A-6	Sequence 6, Appl
15	1051.5	55.5	365	US-08-395-800A-6	Sequence 6, Appl
16	1051.5	55.5	365	US-08-395-800A-10	Sequence 10, Appl
17	1051.5	55.5	365	US-08-696-731-6	Sequence 6, Appl
18	1051.5	55.5	365	US-09-042-531-6	Sequence 6, Appl
19	1051.5	55.5	365	US-09-254-077A-11	Sequence 11, Appl
20	1051.5	55.5	365	PCT-US91-00899-12	Sequence 12, Appl
21	1046.5	55.2	365	US-09-151-592-2	Sequence 2, Appl
22	1046.5	55.2	365	US-09-254-077A-10	Sequence 10, Appl
23	691	36.4	222	US-08-395-800A-2	Sequence 2, Appl
24	439	23.2	102	US-08-393-800A-3	Sequence 3, Appl
25	141	7.4	300	US-09-433-598-2	Sequence 2, Appl
26	97	5.1	183	US-09-167-035-33	Sequence 33, Appl
27	97	5.1	183	US-08-208-887A-33	Sequence 33, Appl

28	97	5.1	183	2	US-08-539-005-33	Sequence 33, Appl
29	97	5.1	183	4	US-09-280-598-35	Sequence 35, Appl
30	96.5	5.1	355	4	US-09-390-131-5	Sequence 5, Appl
31	91.5	4.8	801	1	US-07-906-349A-6	Sequence 6, Appl
32	91	4.8	442	1	US-08-220-151-22	Sequence 22, Appl
33	91	4.8	442	1	US-08-413-118-22	Sequence 22, Appl
34	91	4.8	442	3	US-08-473-446-22	Patent No. 5470718
35	91	4.8	442	6	5470718-3	Sequence 6, Appl
36	87	4.6	217	1	US-08-167-035-6	Sequence 6, Appl
37	87	4.6	217	1	US-08-208-887A-6	Sequence 6, Appl
38	87	4.6	217	2	US-08-539-005-6	Sequence 3, Appl
39	87	4.6	217	2	US-08-815-176-3	Sequence 4, Appl
40	87	4.6	217	2	US-08-815-176-4	Sequence 6, Appl
41	87	4.6	217	4	US-08-664-962B-6	Sequence 6, Appl
42	87	4.6	217	4	US-09-311-743-6	Sequence 6, Appl
43	87	4.6	217	4	US-09-280-598-6	Sequence 6, Appl
44	87	4.6	217	4	US-09-197-344-3	Sequence 3, Appl
45	87	4.6	217	4	US-09-197-344-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-09-298-886-10  
; Sequence 10, Application US/09298886  
; Patent No. 6329170  
; GENERAL INFORMATION:  
; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: 8511-029  
; CURRENT APPLICATION NUMBER: US/09/298,886  
; CURRENT FILING DATE: 1999-04-26  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-298-886-10

Query Match 100.0%; Score 1896; DB 4; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.6e-210;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	LQORIVKLOPLSEKELPMTTQSSGNTESPEMRDSEQHGNGELRGMTTINSIGRLNQM	60
DB	1	LQORIVKLOPLSEKELPMTTQSSGNTESPEMRDSEQHGNGELRGMTTINSIGRLNQM	60
QY	61	GEYATFLALARNGLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER	120
DB	61	GEYATFLALARNGLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER	120
QY	121	YHHPCHVFRVTGYPSCSTFYHHLRPEILKERTLHDHVREREAQAFRLGRVNGSQPSTFV	180
DB	121	YHHPCHVFRVTGYPSCSTFYHHLRPEILKERTLHDHVREREAQAFRLGRVNGSQPSTFV	180
QY	181	GVHVRGDIYVHVPNVKGVADRGYLEKALDMFRARYSSPVFVVTSSNGMWCNENINAS	240
DB	181	GVHVRGDIYVHVPNVKGVADRGYLEKALDMFRARYSSPVFVVTSSNGMWCNENINAS	240
QY	241	RGDVVFAGNIEGSPAKFDALLTQCNHTTMTGTGFWAAYLAGGDTIYLANYTLPDSPF	300
DB	241	RGDVVFAGNIEGSPAKFDALLTQCNHTTMTGTGFWAAYLAGGDTIYLANYTLPDSPF	300
QY	301	LKVFKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFHLKAGVTCYVAGRAF	353
DB	301	LKVFKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFHLKAGVTCYVAGRAF	353

RESULT 2

```

US-09-298-886-8
; Sequence 8, Application US/09298886
; Patent No. 6329170
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/298,886
; CURRENT FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-298-886-8

Query Match      100.0%; Score 1896; DB 4; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.8e-210;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQORIYKQLPLSEKELPMTQSSNGTSEPMRRDSEQHNGELRGMTINSTIGRLGNOM 60
DB 28 LQORIYKQLPLSEKELPMTQSSNGTSEPMRRDSEQHNGELRGMTINSTIGRLGNOM 87
QY 61 GEYATLALARMNGRLAFIPASHMNALAPIFRLSLPVLHSDTAKKIPWNYHLNDWMEER 120
DB 88 GEYATLALARMNGRLAFIPASHMNALAPIFRLSLPVLHSDTAKKIPWNYHLNDWMEER 147
QY 121 YRHIPGHVRFYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSTFV 180
DB 148 YRHIPGHVRFYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSTFV 207
QY 181 GVHVRGDIYVHMPNWKGVADRGYLEKALDMFRARYSPFVVTSGMWCARENINAS 240
DB 208 GVHVRGDIYVHMPNWKGVADRGYLEKALDMFRARYSPFVVTSGMWCARENINAS 257
QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNIHMTGTGFIWAAYLAGGDTIYLANITLPSDF 300
DB 268 RGDVVFAGNGIEGSPAKDFALLTQCNIHMTGTGFIWAAYLAGGDTIYLANITLPSDF 327
QY 301 LKVFKEAAFLPEWVGIPADLSPLK 353
DB 328 LKVFKEAAFLPEWVGIPADLSPLK 380

RESULT 3
US-09-254-077A-6
; Sequence 6, Application US/09254077A
; Patent No. 6399758
; GENERAL INFORMATION:
; APPLICANT: SANDRIN, MAURO S.
; APPLICANT: MCKENZIE, IAN C. F.
; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES
; FILE REFERENCE: 30562.5USWO
; CURRENT APPLICATION NUMBER: US/09/254,077A
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: PCT/AU97/00540
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: PO 1823
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-254-077A-6

Query Match      75.4%; Score 1430; DB 4; Length 340;
Best Local Similarity 81.9%; Pred. No. 1.1e-156;
Matches 267; Conservative 26; Mismatches 19; Indels 14; Gaps 3;

QY 1 LQORIYKQLPLSEKELPMTQSSNGTSEPMRRDSEQHNGELRGMTINSTIGRLGNOM 60
DB 28 LQORMVKIQP--TWELQMTQVT---TESP-----SSPOLKGMWTINAIGRLGNOM 73
QY 61 GEYATLALARMNGRLAFIPASHMNALAPIFRLSLPVLHSDTAKKIPWNYHLNDWMEER 120
DB 74 GEYATLALARMNGRLAFIPASHMNSTLAPIFRITLPLVHASTARRIPWNYHLNDWMEER 133
QY 121 YRHIPGHVRFYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSTFV 180
DB 134 YRHIPGEYVRLTGPCSWTFYHHLRPEILKEFTLHDHVREEAQDFLGLRVNGSRSTFV 193
QY 181 GVHVRGDIYVHMPNWKGVADRGYLEKALDMFRARYSPFVVTSGMWCARENINAS 240
DB 194 GVHVRGDIYVHMPNWKGVADRRYLEQALDWFRARYSPFVVTSGMWCARENINAS 253
QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNIHMTGTGFIWAAYLAGGDTIYLANITLPSDF 300
DB 254 RGDVVFAGNGIEGSPAKDFALLTQCNIHMTGTGFIWAAYLAGGDTIYLANITLPSDF 313

```

```

Matches 267; Conservative 26; Mismatches 19; Indels 14; Gaps 3;

QY 1 LQORIYKQLPLSEKELPMTQSSNGTSEPMRRDSEQHNGELRGMTINSTIGRLGNOM 60
DB 28 LQORMVKIQP--TWELQMTQVT---TESP-----SSPOLKGMWTINAIGRLGNOM 73
QY 61 GEYATLALARMNGRLAFIPASHMNALAPIFRLSLPVLHSDTAKKIPWNYHLNDWMEER 120
DB 74 GEYATLALARMNGRLAFIPASHMNSTLAPIFRITLPLVHASTARRIPWNYHLNDWMEER 133
QY 121 YRHIPGHVRFYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSTFV 180
DB 134 YRHIPGEYVRLTGPCSWTFYHHLRPEILKEFTLHDHVREEAQDFLGLRVNGSRSTFV 193
QY 181 GVHVRGDIYVHMPNWKGVADRGYLEKALDMFRARYSPFVVTSGMWCARENINAS 240
DB 194 GVHVRGDIYVHMPNWKGVADRRYLEQALDWFRARYSPFVVTSGMWCARENINAS 253
QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNIHMTGTGFIWAAYLAGGDTIYLANITLPSDF 300
DB 254 RGDVVFAGNGIEGSPAKDFALLTQCNIHMTGTGFIWAAYLAGGDTIYLANITLPSDF 313

RESULT 4
US-09-254-077A-7
; Sequence 7, Application US/09254077A
; Patent No. 6399758
; GENERAL INFORMATION:
; APPLICANT: SANDRIN, MAURO S.
; APPLICANT: MCKENZIE, IAN C. F.
; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES
; FILE REFERENCE: 30562.5USWO
; CURRENT APPLICATION NUMBER: US/09/254,077A
; CURRENT FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: PCT/AU97/00540
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: PO 1823
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-254-077A-7

Query Match      75.4%; Score 1430; DB 4; Length 340;
Best Local Similarity 81.9%; Pred. No. 1.1e-156;
Matches 267; Conservative 26; Mismatches 19; Indels 14; Gaps 3;

QY 1 LQORIYKQLPLSEKELPMTQSSNGTSEPMRRDSEQHNGELRGMTINSTIGRLGNOM 60
DB 28 LQORMVKIQP--TWELQMTQVT---TESP-----SSPOLKGMWTINAIGRLGNOM 73
QY 61 GEYATLALARMNGRLAFIPASHMNALAPIFRLSLPVLHSDTAKKIPWNYHLNDWMEER 120
DB 74 GEYATLALARMNGRLAFIPASHMNSTLAPIFRITLPLVHASTARRIPWNYHLNDWMEER 133
QY 121 YRHIPGHVRFYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSTFV 180
DB 134 YRHIPGEYVRLTGPCSWTFYHHLRPEILKEFTLHDHVREEAQDFLGLRVNGSRSTFV 193
QY 181 GVHVRGDIYVHMPNWKGVADRGYLEKALDMFRARYSPFVVTSGMWCARENINAS 240
DB 194 GVHVRGDIYVHMPNWKGVADRRYLEQALDWFRARYSPFVVTSGMWCARENINAS 253
QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNIHMTGTGFIWAAYLAGGDTIYLANITLPSDF 300
DB 254 RGDVVFAGNGIEGSPAKDFALLTQCNIHMTGTGFIWAAYLAGGDTIYLANITLPSDF 313

```



```

RESULT 10
PCT-US91-00899-11
: Sequence 11, Application PC/TUS9100899
: GENERAL INFORMATION:
: APPLICANT: Lowe, John B.
: TITLE OF INVENTION: Method and Process
: TITLE OF INVENTION: Oligosaccharide
: TITLE OF INVENTION: or as Free Molecule
: TITLE OF INVENTION: Genetic Sequence
:

```

; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
 ; ADDRESSEE: P.C.  
 ; STREET: 1755 Jefferson Davis Highway, Suite 400  
 ; City: Arlington  
 ; STATE: Virginia  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US91/00899  
 ; FILING DATE: 19910214  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Lavalleye Ph.D., Jean-Paul  
 ; REGISTRATION NUMBER: 31,451  
 ; REFERENCE/DOCKET NUMBER: 2363-021-55 PCT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703)521-5940  
 ; TELEFAX: (703)486-2347  
 ; INFORMATION FOR SEQ ID NO: 11:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 333 amino acids  
 ; TYPE: AMINO ACID  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHEICAL: YES  
 ; FRAGMENT TYPE: C-terminal  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Homo sapiens  
 ; TISSUE TYPE: Blood  
 ; PCT-US91-00899-11

Query Match 55.5%; Score 1051.5; DB 5; Length 333;  
 Best Local Similarity 63.3%; Pred. No. 6.2e-113;  
 Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;  
 QY 25 NTGSPENRRDSEQHNGELRGMTINSIGRLGNMGVATLALARMNGRLAFIPASMH 84  
 Db 27 GTAMGPNASSCPQH-PASLSGTWTVPNGRFGNQMGQYATLLAQLNGRRAFILPAWH 85  
 QY 85 NALAPIRISLPVLHSDTAKKIPWQNYHLNDWMEERYHPIGHFVFTGYPGCSWTFYHHL 144  
 Db 86 AALAPVFRITLFLVLAPEVDSRTPWRLEQLHDWMSSEYADLRDPTLKLSGFPCSWTFHHL 145  
 QY 145 RPEILKEFTLHDHVRREAAQAFRLGLRV--NGSQPSTFVGHVVRGDIYVHVPNQWKGVA 202  
 Db 146 REQIRREFTLHDHLEENAAQSVLGRLGRTGDRPTFVGHVVRGDIYVHVPNQWKGVA 205  
 QY 203 DRGYLEALDMFRARYSSPVFVVTSGNMAWCRENINASRGDVFAGNGIEGSPAKDFALL 262  
 Db 206 DSAYLQAMDMFRARHEAPVFTVTSNGMCKENIDISQGDVTFAGGQZATPKWDFALL 265  
 QY 263 TCQNHITMTGTGFWAAAYLAGGDTIYLANFTLPDSFLKVKFPEAAFLPEWVGIPADLS 322  
 Db 266 TCQNHITMTGTGFWAAAYLAGGDTIYLANFTLPDSFLKVKFPEAAFLPEWVGIPADLS 325  
 QY 323 PLLKALTP 330  
 Db 326 PLWTLAKP 333

## RESULT 11

US-07-914-281-6  
 ; Sequence 6, Application US/07914281  
 ; Patent No. 5324663  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LOWE, JOHN B.

; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
 ; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
 ; TITLE OF INVENTION: GLYCOPOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
 ; ADDRESSEE: P.C.  
 ; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
 ; City: Arlington  
 ; STATE: Virginia  
 ; COUNTRY: U.S.A.  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/914,281  
 ; FILING DATE: 19920720  
 ; CLASSIFICATION: 530  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Lavalleye, Jean-Paul M. P.  
 ; REGISTRATION NUMBER: 31,451  
 ; REFERENCE/DOCKET NUMBER: 2363-060-55  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703)521-4500  
 ; TELEFAX: (703)486-2347  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 365 amino acids  
 ; TYPE: AMINO ACID  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 ; US-07-914-281-6

Query Match 55.5%; Score 1051.5; DB 1; Length 365;  
 Best Local Similarity 63.3%; Pred. No. 7.2e-113;  
 Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;  
 QY 25 NTGSPENRRDSEQHNGELRGMTINSIGRLGNMGVATLALARMNGRLAFIPASMH 84  
 Db 59 GTAMGPNASSCPQH-PASLSGTWTVPNGRFGNQMGQYATLLAQLNGRRAFILPAWH 117  
 QY 85 NALAPIRISLPVLHSDTAKKIPWQNYHLNDWMEERYHPIGHFVFTGYPGCSWTFYHHL 144  
 Db 118 AALAPVFRITLFLVLAPEVDSRTPWRLEQLHDWMSSEYADLRDPTLKLSGFPCSWTFHHL 177  
 QY 145 RPEILKEFTLHDHVRREAAQAFRLGLRV--NGSQPSTFVGHVVRGDIYVHVPNQWKGVA 202  
 Db 178 REQIRREFTLHDHLEENAAQSVLGRLGRTGDRPTFVGHVVRGDIYVHVPNQWKGVA 237  
 QY 203 DRGYLEALDMFRARYSSPVFVVTSGNMAWCRENINASRGDVFAGNGIEGSPAKDFALL 262  
 Db 238 DSAYLQAMDMFRARHEAPVFTVTSNGMCKENIDISQGDVTFAGGQZATPKWDFALL 297  
 QY 263 TCQNHITMTGTGFWAAAYLAGGDTIYLANFTLPDSFLKVKFPEAAFLPEWVGIPADLS 322  
 Db 298 TCQNHITMTGTGFWAAAYLAGGDTIYLANFTLPDSFLKVKFPEAAFLPEWVGIPADLS 357  
 QY 323 PLLKALTP 330  
 Db 358 PLWTLAKP 365

## RESULT 12

US-08-393-246-6  
 ; Sequence 6, Application US/08393246  
 ; Patent No. 5595900  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LOWE, JOHN B.

;; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
;; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
;; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
;; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
;; NUMBER OF SEQUENCES: 14  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
;; ADDRESSEE: P.C.  
;; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
;; CITY: Arlington  
;; STATE: Virginia  
;; COUNTRY: U.S.A.  
;; ZIP: 22202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/393,246  
;; FILING DATE:  
;; CLASSIFICATION: 530  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/220,433  
;; FILING DATE: 30-MAR-1994  
;; APPLICATION NUMBER: US 07/914,281  
;; FILING DATE: 20-JUL-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Lavalleye, Jean-Paul M. P.  
;; REGISTRATION NUMBER: 31,451  
;; REFERENCE/DOCKET NUMBER: 2363-060-55  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703)521-4500  
;; TELEFAX: (703)486-2347  
;; TELEX: 248855 OPAT UR  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 365 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: protein  
;; US-08-393-246-6

Query Match 55.5%; Score 1051.5; DB 1; Length 365;  
Best Local Similarity 63.3%; Pred. No. 7.2e-113;  
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;  
QY 25 GNTSPENRRDSEHGNGELRGMTFTINSIGRLGNQMGEXATLFLALRMNGRLAFIPASMH 84  
Db 59 GTAMGPNASSSCPOH-PASLSGIWTVPNGRFGNMGQYATLLAALQNGRRAPILPAMH 117  
QY 85 NALAPIFRISLPVLHSDTAKKIPQNYHLNDWMBEERYHPIGHEVFTGYPGCSWTFYHHL 144  
Db 118 AALAPVFRITLPLVAPEVDSRTPWRELQHDWSEEVADLRDPFLKSGFPCSWTFHHL 177  
QY 145 RPEILKEFTLHDHVREAAQAFRLGLRV--NGSQPSTFVGHVHVRGDYVHYMNPVWGKVA 202  
Db 178 REQIRREFTLHDHREAAQSVLGLRLGRTGDRPRTFVGHVHVRGDYVHYMNPVWGKVA 237  
QY 203 DRGYLEKALDMFRARYSSPVFTVTSNGMAMRENINASRGDVPVFNAGNIGSPAKDFALL 262  
Db 238 DSAYLRQAMDWFRARHEAPVFEVVTISNGMCKENIDTSQGDVTFAGDQCATPWKDFALL 297  
QY 263 TCQNTHTMTTGTGWAAYLAGDVTIYLANYTLPDPSFLKVFKEPAFLPEWVGIPADLS 322  
Db 298 TCQNTHTMTTGTGWAAYLAGDVTIYLANYTLPDPSFLKVFKEPAFLPEWVGIPADLS 357  
QY 323 PLLKALIP 330  
Db 358 PLWTAKP 365

RESULT 13

US-08-273-411-1  
;; Sequence 1, Application US/08273411  
;; Patent No. 5625124  
;; GENERAL INFORMATION:  
;; APPLICANT: Falk, Per  
;; APPLICANT: Gordon, Jeffrey I.  
;; TITLE OF INVENTION: Animal Model for Gastro-Intestinal  
;; TITLE OF INVENTION: Disease  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Patrea L. Pabst  
;; STREET: 1100 Peachtree Street, Suite 2800  
;; CITY: Atlanta  
;; STATE: Georgia  
;; COUNTRY: USA  
;; ZIP: 30309-4530  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/273,411  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Pabst, Patrea L.  
;; REGISTRATION NUMBER: 31,284  
;; REFERENCE/DOCKET NUMBER: WU106  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (404) 815-6508  
;; TELEFAX: (404) 815-6555  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 365 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOTHEICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: 1..365  
;; OTHER INFORMATION: /note= "GDP-L-fucose:beta-D-Galactoside-2-alpha-L-fucosyl  
;; PUBLICATION INFORMATION:  
;; AUTHORS: Larsen, et al.  
;; JOURNAL: Proc. Nat'l Acad. Sci. USA  
;; VOLUME: 87  
;; PAGES: 6674-6678  
;; DATE: 1990  
;; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 365  
US-08-273-411-1

Query Match 55.5%; Score 1051.5; DB 1; Length 365;  
Best Local Similarity 63.3%; Pred. No. 7.2e-113;  
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;  
QY 25 GNTSPENRRDSEHGNGELRGMTFTINSIGRLGNQMGEXATLFLALRMNGRLAFIPASMH 84  
Db 59 GTAMGPNASSSCPOH-PASLSGIWTVPNGRFGNMGQYATLLAALQNGRRAPILPAMH 117  
QY 85 NALAPIFRISLPVLHSDTAKKIPQNYHLNDWMBEERYHPIGHEVFTGYPGCSWTFYHHL 144  
Db 118 AALAPVFRITLPLVAPEVDSRTPWRELQHDWSEEVADLRDPFLKSGFPCSWTFHHL 177  
QY 145 RPEILKEFTLHDHVREAAQAFRLGLRV--NGSQPSTFVGHVHVRGDYVHYMNPVWGKVA 202  
Db 178 REQIRREFTLHDHREAAQSVLGLRLGRTGDRPRTFVGHVHVRGDYVHYMNPVWGKVA 237  
QY 203 DRGYLEKALDMFRARYSSPVFTVTSNGMAMRENINASRGDVPVFNAGNIGSPAKDFALL 262

Db 238 DSAYLRQAMDFRAREAPVFFVVTNSGMWCKENIDTSQGDVTFAGDGOEATPWKDFALL 297

Qy 263 TOCNHTMTIGTGWAAYLGGDDTIYLANITLPSDFLKVKFPEAAFLPEWVGIPADLS 322

Db 298 TOCNHTMTIGTGWAAYLGGDDTIYLANITLPSDFLKVKFPEAAFLPEWVGIPADLS 357

Qy 323 PLLKALTP 330

Db 358 PLWTLAKP 365

RESULT 14

US-08-525-058A-6

Sequence 6, Application US/08525058A

Patent No. 5770420

GENERAL INFORMATION:

APPLICANT: LOWE, JOHN B.

TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS

TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS.

TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION

TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.

STREET: 1755 Jefferson Davis Highway, Fourth Floor

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/525,058A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Lavalleye, Jean-Paul M. P.

REGISTRATION NUMBER: 31,451

REFERENCE/DOCKET NUMBER: 2363-060-55

TELEPHONE: (703)521-4500

TELEFAX: (703)486-2347

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 365 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-525-058A-6

Query Match 55.5%; Score 1051.5; DB 1; Length 365;

Best Local Similarity 63.3%; Pred. No. 7.2e-113;

Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

Qy 25 GNTSPKMRDSEOHNGELGMFTINSIGRLGNMGYATILFALARMNGRIAFIPASMH 84

Db 59 GTAMGNASSCPQH-PASLSGTWTVPYNGRGNMGYATILFALARMNGRIAFIPASMH 117

Qy 85 NALAPFIRISLPVLHSDTAKKIPWNYHLNDWMEERYRHIPGHFVFRFTGFCSTFFYHHL 144

Db 118 AALAPVFRTITPLVAPEVDSRTPWRELQHDWMSSEYADLRDPFLKLSGFFCSTFFYHHL 177

Qy 145 RPEILKFTLHDHVEEAQAQLRLGV--NGSQPTFFGVHVRGDYVHVMPNWKGVVA 202

Db 178 REQIRREFTLHDHLEEAQAQLRLGLRGTDRPRTFVGHVHVRGDYLVQVMPQKWKGVV 237

Qy 203 DRGYLEKALDMFRARYSSPVFVVTNSGMWCKENINASRGDVFVFNAGTSGSPAKDFALL 262

Db 238 DSAYLRQAMDFRAREAPVFFVVTNSGMWCKENIDTSQGDVTFAGDGOEATPWKDFALL 297

Qy 263 TOCNHTMTIGTGWAAYLGGDDTIYLANITLPSDFLKVKFPEAAFLPEWVGIPADLS 322

Db 298 TOCNHTMTIGTGWAAYLGGDDTIYLANITLPSDFLKVKFPEAAFLPEWVGIPADLS 357

Qy 323 PLLKALTP 330

Db 358 PLWTLAKP 365

RESULT 15

US-08-395-800A-6

Sequence 6, Application US/08395800A

Patent No. 5807732

GENERAL INFORMATION:

APPLICANT: LOWE, JOHN B

APPLICANT: LENNON, GREGORY

APPLICANT: ROQUER, SYLVIE

APPLICANT: GIORGI, DOMINIQUE

APPLICANT: KELLY, ROBERT J

TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE

TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE

TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESSES:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT

STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/395,800A

FILING DATE: 28-FEB-1995

CLASSIFICATION: 435

TELEPHONE: (703) 413-3000

TELEFAX: (703) 413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 365 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-395-800A-6

Query Match 55.5%; Score 1051.5; DB 1; Length 365;

Best Local Similarity 63.3%; Pred. No. 7.2e-113;

Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

Qy 25 GNTSPKMRDSEOHNGELGMFTINSIGRLGNMGYATILFALARMNGRIAFIPASMH 84

Db 59 GTAMGNASSCPQH-PASLSGTWTVPYNGRGNMGYATILFALARMNGRIAFIPASMH 117

Qy 85 NALAPFIRISLPVLHSDTAKKIPWNYHLNDWMEERYRHIPGHFVFRFTGFCSTFFYHHL 144

Db 118 AALAPVFRTITPLVAPEVDSRTPWRELQHDWMSSEYADLRDPFLKLSGFFCSTFFYHHL 177

Qy 145 RPEILKFTLHDHVEEAQAQLRLGV--NGSQPTFFGVHVRGDYVHVMPNWKGVVA 202

Db 178 REQIRREFTLHDHLEEAQAQLRLGLRGTDRPRTFVGHVHVRGDYLVQVMPQKWKGVV 237

Qy 203 DRGYLEKALDMFRARYSSPVFVVTNSGMWCKENINASRGDVFVFNAGTSGSPAKDFALL 262

Db 238 DSAYLRQAMDFRAREAPVFFVVTNSGMWCKENIDTSQGDVTFAGDGOEATPWKDFALL 297

us-10-040-863-10.rai

Wed May 28 09:21:51 2003

Qy 263 TOCNHTIMTIGTIGIWAAYLAGGDTIYLANYTLPDSPFLKYFKPEAFLEPWGIPADLS 322  
Db 298 TOCNHTIMTIGTIGFGEWAAYLAGGDTVYLANETLPDSEELKIFKPEAFLEPWGINADLS 357  
Qy 323 PLLKALTP 330  
Db 358 PLWTLAKP 365

Search completed: May 27, 2003, 15:11:52  
Job time : 15.5136 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:08:29 ; Search time 15.2516 Seconds  
(without alignments)  
2295.367 Million cell updates/sec

Title: US-10-040-863-10

Perfect score: 1896

Sequence: 1 LQQRIVKLPISKEPLMTT.....RSHFLKAKGYCYVAGRAF 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1896	100.0	353	10	US-09-999-672-10
2	1896	100.0	353	12	US-10-040-863-10
3	1896	100.0	380	10	US-09-999-672-8
4	1896	100.0	380	12	US-10-040-863-8
5	1430	75.4	340	10	US-09-051-034A-2
6	1376	72.6	344	10	US-09-999-672-11
7	1376	72.6	344	12	US-10-040-863-11
8	1051.5	55.5	365	9	US-10-105-963-10
9	1051.5	55.5	365	10	US-09-863-475A-6
10	1046.5	55.2	365	10	US-09-051-034A-4
11	1044.5	55.1	365	10	US-09-844-268-13
12	1044.5	55.1	365	10	US-09-844-705-13
13	141	7.4	300	10	US-09-848-838-2
14	99	5.2	436	9	US-10-214-524-25
15	97.5	5.1	431	9	US-09-479-614-14
16	97.5	5.1	436	9	US-09-479-614-2
17	97.5	5.1	496	9	US-09-764-614-29
18	89	4.7	1031	9	US-09-764-868-647
19	89	4.7	2057	9	US-09-815-379-10

```
89 4.7 2058 9 US-09-815-379-17 Sequence 17, Appli
21 87 217 10 US-09-765-298A-6 Sequence 6, Appli
22 85.5 738 9 US-10-235-521-1 Sequence 1, Appli
23 84.5 344 9 US-09-895-913A-56 Sequence 56, Appli
24 84.5 344 10 US-09-815-242-11366 Sequence 11366, A
25 83.5 747 9 US-10-028-072-212 Sequence 212, App
26 83.5 747 9 US-10-121-049-212 Sequence 212, App
27 83.5 747 9 US-10-123-904-212 Sequence 212, App
28 83.5 747 9 US-10-140-470-212 Sequence 212, App
29 83.5 747 9 US-10-175-746-212 Sequence 212, App
30 83.5 747 9 US-10-176-918-212 Sequence 212, App
31 83.5 747 9 US-10-176-921-212 Sequence 212, App
32 83.5 747 9 US-10-137-865-212 Sequence 212, App
33 83.5 747 9 US-10-140-474-212 Sequence 212, App
34 83.5 747 9 US-10-142-431-212 Sequence 212, App
35 83.5 747 9 US-10-143-114-212 Sequence 212, App
36 83.5 747 9 US-10-140-002-212 Sequence 212, App
37 83.5 747 9 US-10-142-419-212 Sequence 212, App
38 83.5 747 9 US-10-123-263-212 Sequence 212, App
39 83.5 747 9 US-10-142-423-212 Sequence 212, App
40 83.5 747 9 US-10-121-050-212 Sequence 212, App
41 83.5 747 9 US-10-141-753-212 Sequence 212, App
42 83.5 747 9 US-10-143-032-212 Sequence 212, App
43 83.5 747 9 US-10-123-108-212 Sequence 212, App
44 83.5 747 9 US-10-123-236-212 Sequence 212, App
45 83.5 747 9 US-10-123-261-212 Sequence 212, App
```

#### ALIGNMENTS

```
RESULT 1
US-09-999-672-10
; Sequence 10, Application US/09999672
; Patent No. US20020127655A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/999,672
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US/09/298,886
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-999-672-10
```

```
Query Match 100.0%; Score 1896; DB 10; Length 353;
Best Local Similarity 100.0%; Pred. No. 3.9e-188;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQQRIVKLPISKEPLMTTQSSNGTSPSPMRDRSEOHNGELRGMTINSIGLGNOM 60
|||||
Db 1 LQQRIVKLPISKEPLMTTQSSNGTSPSPMRDRSEOHNGELRGMTINSIGLGNOM 60
|||||

QY 61 GEYATLALARMNRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWKEER 120
|||||
Db 61 GEYATLALARMNRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWKEER 120
|||||

QY 121 YRHLPGHVFRTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGRVNGSQPSTFV 180
|||||
Db 121 YRHLPGHVFRTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGRVNGSQPSTFV 180
|||||

QY 181 GVHVRGDDYVHVMNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRENINAS 240
|||||
Db 181 GVHVRGDDYVHVMNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRENINAS 240
|||||
```

```
QY 241 RGVVFAGNIEGSPAKDFALLTQCNTTMTIGTFGIWAAYLAGGDTIYLANYTLPDSPF 300
D 241 RGVVFAGNIEGSPAKDFALLTQCNTTMTIGTFGIWAAYLAGGDTIYLANYTLPDSPF 300
QY 301 LKVKPEAAFLPEWVGIPADLSPLLKALTACPRSHFLKAKGVTCYVAGRAF 353
D 301 LKVKPEAAFLPEWVGIPADLSPLLKALTACPRSHFLKAKGVTCYVAGRAF 353

RESULT 2
US-10-040-863-10
; Sequence 10, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/298,886
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-040-863-10

Query Match 100.0%; Score 1896; DB 12; Length 353;
Best Local Similarity 100.0%; Pred. No. 3.9e-188;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQORIVKLOPLSEKELPMTQSSNGTSPENRRDSEQHNGELRGMTINSIGRLGNQM 60
D 1 LQORIVKLOPLSEKELPMTQSSNGTSPENRRDSEQHNGELRGMTINSIGRLGNQM 60
QY 61 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
D 61 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
QY 121 YRHIPGHFVRETGYPCSWTFYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSQPSTFV 180
D 121 YRHIPGHFVRETGYPCSWTFYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSQPSTFV 180
QY 121 YRHIPGHFVRETGYPCSWTFYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSQPSTFV 180
D 121 YRHIPGHFVRETGYPCSWTFYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSQPSTFV 180
QY 181 GVHVRGDIYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCARENINAS 240
D 181 GVHVRGDIYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCARENINAS 240
QY 241 RGVVFAGNIEGSPAKDFALLTQCNTTMTIGTFGIWAAYLAGGDTIYLANYTLPDSPF 300
D 241 RGVVFAGNIEGSPAKDFALLTQCNTTMTIGTFGIWAAYLAGGDTIYLANYTLPDSPF 300
QY 301 LKVKPEAAFLPEWVGIPADLSPLLKALTACPRSHFLKAKGVTCYVAGRAF 353
D 301 LKVKPEAAFLPEWVGIPADLSPLLKALTACPRSHFLKAKGVTCYVAGRAF 353

RESULT 3
US-09-999-672-8
; Sequence 8, Application US/09999672
; Patent No. US2002012765A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/999,672
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US/09/298,886
```

```
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-999-672-8

Query Match 100.0%; Score 1896; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 4.4e-188;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQORIVKLOPLSEKELPMTQSSNGTSPENRRDSEQHNGELRGMTINSIGRLGNQM 60
D 28 LQORIVKLOPLSEKELPMTQSSNGTSPENRRDSEQHNGELRGMTINSIGRLGNQM 87
QY 61 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
D 83 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER 147
QY 121 YRHIPGHFVRETGYPCSWTFYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSQPSTFV 180
D 148 YRHIPGHFVRETGYPCSWTFYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSQPSTFV 207
QY 181 GVHVRGDIYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCARENINAS 240
D 208 GVHVRGDIYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCARENINAS 267
QY 241 RGVVFAGNIEGSPAKDFALLTQCNTTMTIGTFGIWAAYLAGGDTIYLANYTLPDSPF 300
D 268 RGVVFAGNIEGSPAKDFALLTQCNTTMTIGTFGIWAAYLAGGDTIYLANYTLPDSPF 327
QY 301 LKVKPEAAFLPEWVGIPADLSPLLKALTACPRSHFLKAKGVTCYVAGRAF 353
D 328 LKVKPEAAFLPEWVGIPADLSPLLKALTACPRSHFLKAKGVTCYVAGRAF 380

RESULT 4
US-10-040-863-8
; Sequence 8, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/298,886
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-040-863-8

Query Match 100.0%; Score 1896; DB 12; Length 380;
Best Local Similarity 100.0%; Pred. No. 4.4e-188;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQORIVKLOPLSEKELPMTQSSNGTSPENRRDSEQHNGELRGMTINSIGRLGNQM 60
D 28 LQORIVKLOPLSEKELPMTQSSNGTSPENRRDSEQHNGELRGMTINSIGRLGNQM 87
QY 61 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
D 88 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER 147
QY 121 YRHIPGHFVRETGYPCSWTFYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSQPSTFV 180
```

DB 148 YRHIFGHEVFRFTGYPSCWTFYHHLRPEILKEFTLHDHVRERAAQFLRGLRVNGSQSPSFV 207  
QY 181 GVHVRGDIYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCNENAS 240  
DB 208 GVHVRGDIYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCNENAS 267  
QY 241 RGVVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLANITLPDPSF 300  
DB 268 RGVVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLANITLPDPSF 327  
QY 301 LKVFPEAAFLPEWVGIPADLSPLKALTPACPRSHFLKAGVTCYVAGRAF 353  
DB 328 LKVFPEAAFLPEWVGIPADLSPLKALTPACPRSHFLKAGVTCYVAGRAF 380

## RESULT 5

US-09-051-034A-2

; Sequence 2, Application US/09051034A

; Patent No. US2001005584A1

; GENERAL INFORMATION:

; APPLICANT: MCKENZIE, IAN FAROUHAR CAMPBELL

; APPLICANT: SANDRIN, MAURO SERGIO

; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC

; TITLE OF INVENTION: GLYCOSYLTRANSFERASE

; FILE REFERENCE: 30562.60SWO

; CURRENT APPLICATION NUMBER: US/09/051.034A

; PRIOR FILING DATE: 1998-03-31

; PRIOR FILING DATE: 1997-08-01

; PRIOR FILING DATE: 1996-08-21

; PRIOR FILING DATE: 1996-08-02

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 340

; TYPE: PRT

; ORGANISM: Sus Domesticus

US-09-051-034A-2

Query Match 75.4%; Score 1430; DB 10; Length 340;  
Best Local Similarity 81.9%; Pred. No. 7.7e-140;  
Matches 267; Conservative 26; Mismatches 19; Indels 14; Gaps 3;

QY 1 LQORIVKLOPLSEKELPMTTOMSSGNTESPERRDSEOHNGELRGMTINSIGRLGNQM 60  
DB 28 LQORIVKLOPLSEKELPMTTOMSSGNTESPERRDSEOHNGELRGMTINSIGRLGNQM 73  
QY 61 GEYATLALARMNGRLAFIPASMHNALAPIERISLPVLHSDTAKKIPWONYHLNDWMEER 120  
DB 74 GEYATLALARMNGRLAFIPASMHNALAPIERISLPVLHSDTAKKIPWONYHLNDWMEER 133  
QY 121 YRHIFGHEVFRFTGYPSCWTFYHHLRPEILKEFTLHDHVRERAAQFLRGLRVNGSQSPSFV 180  
DB 134 YRHIFGHEVFRFTGYPSCWTFYHHLRPEILKEFTLHDHVRERAAQFLRGLRVNGSQSPSFV 193  
QY 181 GVHVRGDIYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCNENAS 240  
DB 194 GVHVRGDIYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCNENAS 253  
QY 241 RGVVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLANITLPDPSF 300  
DB 254 RGVVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLANITLPDPSF 313  
QY 301 LKVFPEAAFLPEWVGIPADLSPLK 326  
DB 314 LKVFPEAAFLPEWVGIPADLSPLK 339

## RESULT 6

US-09-999-672-11

; Sequence 2, Application US/09051034A

; Patent No. US2001005584A1

; GENERAL INFORMATION:

; APPLICANT: MCKENZIE, IAN FAROUHAR CAMPBELL

; APPLICANT: SANDRIN, MAURO SERGIO

; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC

; TITLE OF INVENTION: GLYCOSYLTRANSFERASE

; FILE REFERENCE: 30562.60SWO

; CURRENT APPLICATION NUMBER: US/09/051.034A

; PRIOR FILING DATE: 1998-03-31

; PRIOR FILING DATE: 1997-08-01

; PRIOR FILING DATE: 1996-08-21

; PRIOR FILING DATE: 1996-08-02

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 340

; TYPE: PRT

; ORGANISM: Sus Domesticus

US-09-051-034A-2

; Sequence 11, Application US/09999672  
; Patent No. US20020127655A1  
; GENERAL INFORMATION:  
; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: 8511-029  
; CURRENT APPLICATION NUMBER: US/09/999,672  
; CURRENT FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: US/09/298,886  
; PRIOR FILING DATE: 1999-04-26  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 344  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-999-672-11

Query Match 72.6%; Score 1376; DB 10; Length 344;  
Best Local Similarity 78.9%; Pred. No. 3.1e-134;  
Matches 258; Conservative 26; Mismatches 31; Indels 12; Gaps 3;

QY 1 LQORIVKLOPLSEKELPMTTOMSSGNTESPERRDSEOHNGELRGMTINSIGRLGNQM 60  
DB 28 VQORLAKIQAM--WELPY-----QIPVLASTSKALGPSQLRGWMTINAIGRLGNQM 76  
QY 61 GEYATLALARMNGRLAFIPASMHNALAPIERISLPVLHSDTAKKIPWONYHLNDWMEER 120  
DB 77 GEYATLALARMNGRLAFIPASMHNALAPIERISLPVLHSDTAKKIPWONYHLNDWMEER 136  
QY 121 YRHIFGHEVFRFTGYPSCWTFYHHLRPEILKEFTLHDHVRERAAQFLRGLRVNGSQSPSF 179  
DB 137 YRHIFGHEVFRFTGYPSCWTFYHHLRPEILKEFTLHDHVRERAAQFLRGLRVNGSQSPSF 196  
QY 180 GVHVRGDIYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCNENAS 239  
DB 197 GVHVRGDIYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCNENAS 256  
QY 240 RGVVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLANITLPDPSF 299  
DB 257 RGVVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLANITLPDPSF 316  
QY 300 LKVFPEAAFLPEWVGIPADLSPLK 326  
DB 317 LKVFPEAAFLPEWVGIPADLSPLK 343

## RESULT 7

US-10-040-863-11

; Sequence 11, Application US/10040863

; Patent No. US20020137165A1

; GENERAL INFORMATION:

; APPLICANT: Eric H. Holmes et al.

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE

; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES

; FILE REFERENCE: 8511-029

; CURRENT APPLICATION NUMBER: US/10/040,863

; PRIOR APPLICATION NUMBER: 09/298,886

; PRIOR FILING DATE: 1999-04-23

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 11

; LENGTH: 344

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-040-863-11

Query Match 72.6%; Score 1376; DB 12; Length 344;  
Best Local Similarity 78.9%; Pred. No. 3.1e-134;

```
Matches 258; Conservative 26; Mismatches 31; Indels 12; Gaps 3;
QY 1 LQORTVKLOPLSEKELPMTOMSGNTSEPMRDSQHGNGELRGMTINSIGRLGNQM 60
DB 28 VQORLAKTQAM--WELPV-----QIPVLASTSKALGPSQLRGMTINAGRLGNQM 76
QY 61 GEYATFLALAMNGRLATPASHMNALAPIRISLPLVHSDTAKKIPWNYHLNDWMEER 120
DB 77 GEYATFLALAMNGRLATPASHMNALAPIRISLPLVHSDTAKKIPWNYHLNDWMEER 136
QY 121 YRHI-PGHFVFTGYPGKSWTFYHHLRPEILKEFTLHDHVREDAQAFRLGRVNSQPTF 179
DB 137 YHIPPGEYVRETGYPGKSWTFYHHLRPEILKEFTLHDHVREDAQAFRLGRVNSQPTF 196
QY 180 VGVHVRGDDYVHVPNNVKNVGVVADRGYLEKALDMPFRARYSSPVFVTSNGMAWCHRENA 239
DB 197 VGVHVRGDDYVHVPNNVKNVGVVADRGYLEKALDMPFRARYSSPVFVTSNGMAWCHRENA 256
QY 240 SRGDVVFAGDGGIEGSPAKDFALLTCNHTINTIGTGFIAWAAVLGAGDIIYLANIYLPDSP 299
DB 257 SHGDVVFAGDGGIEGSPAKDFALLTCNHTINTIGTGFIAWAAVLGAGDIIYLANIYLPDSP 316
QY 300 FLKVKPEAAFLPEWVGIPADLSPLLK 326
DB 317 FLKVKPEAAFLPEWVGIPADLSPLLK 343
RESULT 8
US-10-105-963-10
; Sequence 10, Application US/10105963
; Publication No. US20030068818A1
; GENERAL INFORMATION:
; APPLICANT: Genon Corporation
; APPLICANT: Denning, Chris
; APPLICANT: Clark, A. John
; APPLICANT: Schliff, J. Michael
; TITLE OF INVENTION: Animal Tissue with Carbohydrate Antigens Compatible for Human
; TITLE OF INVENTION: Transplantation and a Carbohydrate Determinant Selection System
; FILE REFERENCE: 731/002
; CURRENT APPLICATION NUMBER: US/10/105,963
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: US 60/277,811
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1.0
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-105-963-10
Query Match 55.5%; Score 1051.5; DB 9; Length 365;
Best Local Similarity 63.3%; Pred. No. 1.5e-100;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 25 GNTSEPMRDSQHGNGELRGMTINSIGRLGNOMGEYATFLALAMNGRLAFIPASHM 84
DB 59 GTAMGNASSCCPQH-PASLSGTWTVPNGRFGNQMGQYATLLALAGLNRRRAFLPAMH 117
QY 85 NALAPIRISLPLVHSDTAKKIPWNYHLNDWMEERYRHIPGHFVFTGYPGKSWTFYHHL 144
DB 118 AALAPVFRITLPLVLAPEVDSRTWRELQHDWMESEYADLRDPFLKLSGFCPSWTFPHL 177
QY 145 RPEILKEFTLHDHVREDAQAFRLGRV--NGSQPSFTGVHVRGDDYVHVPNNVKNVGVV 202
DB 178 REQIRREFTLHDHLREDAQSVGLQRLGRGDRPRTGVHVRGDDYVHVPNNVKNVGVV 237
QY 203 DRGVLEKALDMFRARYSSPVFVTSNGMAWCHRENAINASRGDVVFAGDIEGSPAKDFALL 262
DB 238 DSAYLRQAMDWRARHEAPVFTVTSNGMAWCHRENAINASRGDVVFAGDIEGSPAKDFALL 297
QY 263 TQCNHTINTIGTGFIAWAAVLGAGDIIYLANIYLPDPSPLLKVKPEAAFLPEWVGIPADLS 322
```

```
DB 298 TQCNHTINTIGTGFIAWAAVLGAGDIIYLANIYLPDPSPLLKVKPEAAFLPEWVGIPADLS 357
QY 323 PLLKALTP 330
DB 358 PLWTLAKP 365
RESULT 9
US-09-863-475A-6
; Sequence 6, Application US/09863475A
; Patent No. US20020102688A1
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
City: Arlington
State: Virginia
Country: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/863,475A
FILING DATE: 24-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-863-475A-6
Query Match 55.5%; Score 1051.5; DB 10; Length 365;
Best Local Similarity 63.3%; Pred. No. 1.5e-100;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 25 GNTSEPMRDSQHGNGELRGMTINSIGRLGNOMGEYATFLALAMNGRLAFIPASHM 84
DB 59 GTAMGNASSCCPQH-PASLSGTWTVPNGRFGNQMGQYATLLALAGLNRRRAFLPAMH 117
QY 85 NALAPIRISLPLVHSDTAKKIPWNYHLNDWMEERYRHIPGHFVFTGYPGKSWTFYHHL 144
DB 118 AALAPVFRITLPLVLAPEVDSRTWRELQHDWMESEYADLRDPFLKLSGFCPSWTFPHL 177
QY 145 RPEILKEFTLHDHVREDAQAFRLGRV--NGSQPSFTGVHVRGDDYVHVPNNVKNVGVV 202
DB 178 REQIRREFTLHDHLREDAQSVGLQRLGRGDRPRTGVHVRGDDYVHVPNNVKNVGVV 237
QY 203 DRGVLEKALDMFRARYSSPVFVTSNGMAWCHRENAINASRGDVVFAGDIEGSPAKDFALL 262
```

Db 238 DSAYLRQDMFRARHEAPVFTVTSNGMEWCKENIDTSQGVTFAGDQGEATPKRDFALL 297  
Qy 263 TQCNTIMTIGTGIWAAYLAGGDTIYLANITLPSPLKVKFKEAAFLPEWVGIPADLS 322  
Db 298 TQCNTIMTIGTGFWAAYLAGGDTIYLANITLPSPLKVKFKEAAFLPEWVGINADLS 357  
Qy 323 PLLKALTP 330  
Db 358 PLWTIAKP 365

## RESULT 10

US-09-051-034A-4  
; Sequence 4, Application US/09051034A  
; Patent No. US20010055384A1  
; GENERAL INFORMATION:  
; APPLICANT: MCKENZIE, IAN FAROUHAR CAMPBELL  
; APPLICANT: SANDRIN, MAURO SERGIO  
; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC  
; TITLE OF INVENTION: GLYCOSYLTRANSFERASE  
; FILE REFERENCE: 30562.6USWO  
; CURRENT APPLICATION NUMBER: US/09/051.034A  
; CURRENT FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: PCT/AU97/00492  
; PRIOR FILING DATE: 1997-08-01  
; PRIOR APPLICATION NUMBER: 60/024,279  
; PRIOR FILING DATE: 1996-08-21  
; PRIOR APPLICATION NUMBER: P01402  
; PRIOR FILING DATE: 1996-08-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 365  
; TYPE: PRT  
; ORGANISM: Sus Domesticus  
US-09-051-034A-4

Query Match 55.2%; Score 1046.5; DB 10; Length 365;  
Best Local Similarity 62.6%; Pred. No. 5e-100;  
Matches 194; Conservative 38; Mismatches 75; Indels 3; Gaps 2;

Qy 17 PMTQMSSGNTESPEMRDSEHGNGELRGMTFTNSIGRLGNQMGGEYATLALARMNGRL 76  
Db 51 PVALIFCLAGTPVHPNADSCPKH-PASFGTWTIYDPGRFGNQMGQYATLLALAGLNGRQ 109  
Qy 77 AFIPASMNALAPIRISLPVLHSDTAKKIPQNYHLNDMMEERYRHPGHVFRFTGYPC 136  
Db 110 AFIQPAMHVLAPVFRITLPVLAPEVDHAPWRELELDHWMSEDYAHLKEPWLKLTGPPC 169  
Qy 137 SWTFVYHLRPEILKEFTLHDHVRREAQAFRLGLRV--NGSOPSTFVGHVVRGDDYVHVMP 194  
Db 170 SWTFVYHLRPEILKEFTLHDHVRREAQAFRLGLRV--NGSOPSTFVGHVVRGDDYVHVMP 229  
Qy 195 NWKGVVADRGYLEKALDMFRARYSSPVFTVTSNGMACRENINASRGDVFVAGNGIEGS 254  
Db 230 KRWKGVVGDGYLQAMDWFARYEAPVFTVTSNGMEWCKENIDTSRGDVFVAGNGIEGS 289  
Qy 255 PAKDFALLTQCNHMTIGTGFVWAAYLAGGDTIYLANITLPSPLKVKFKEAAFLPEW 314  
Db 290 PARDFALLVQCNHMTIGTGFVWAAYLAGGDTIYLANITLPSPLKVKFKEAAFLPEW 349  
Qy 315 VGIPADLSPL 324  
Db 350 VGINADLSPL 359

## RESULT 11

US-09-844-268-13  
; Sequence 13, Application US/09844268  
; Patent No. US20020129395A1  
; GENERAL INFORMATION:  
; APPLICANT: BOSWORTH, BRAD

; APPLICANT: VOGELI, PETER  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY  
; TITLE OF INVENTION: RESISTANT TO F18 E. COLI ASSOCIATED DISEASES  
; FILE REFERENCE: 21419/90368  
; CURRENT APPLICATION NUMBER: US/09/844,268  
; CURRENT FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: 09/443,766  
; PRIOR FILING DATE: 1999-11-19  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 365  
; TYPE: PRT  
; ORGANISM: Porcine  
US-09-844-268-13

Query Match 55.1%; Score 1044.5; DB 10; Length 365;  
Best Local Similarity 62.6%; Pred. No. 8e-100;  
Matches 194; Conservative 38; Mismatches 75; Indels 3; Gaps 2;

Qy 17 PMTQMSSGNTESPEMRDSEHGNGELRGMTFTNSIGRLGNQMGGEYATLALARMNGRL 76  
Db 51 PVALIFCLAGTPVHPNADSCPKH-PASFGTWTIYDPGRFGNQMGQYATLLALAGLNGRQ 109  
Qy 77 AFIPASMNALAPIRISLPVLHSDTAKKIPQNYHLNDMMEERYRHPGHVFRFTGYPC 136  
Db 110 AFIQPAMHVLAPVFRITLPVLAPEVDHAPWRELELDHWMSEDYAHLKEPWLKLTGPPC 169  
Qy 137 SWTFVYHLRPEILKEFTLHDHVRREAQAFRLGLRV--NGSOPSTFVGHVVRGDDYVHVMP 194  
Db 170 SWTFVYHLRPEILKEFTLHDHVRREAQAFRLGLRV--NGSOPSTFVGHVVRGDDYVHVMP 229  
Qy 195 NWKGVVADRGYLEKALDMFRARYSSPVFTVTSNGMACRENINASRGDVFVAGNGIEGS 254  
Db 230 KRWKGVVGDGYLQAMDWFARYEAPVFTVTSNGMEWCKENIDTSRGDVFVAGNGIEGS 289  
Qy 255 PAKDFALLTQCNHMTIGTGFVWAAYLAGGDTIYLANITLPSPLKVKFKEAAFLPEW 314  
Db 290 PARDFALLVQCNHMTIGTGFVWAAYLAGGDTIYLANITLPSPLKVKFKEAAFLPEW 349  
Qy 315 VGIPADLSPL 324  
Db 350 VGINADLSPL 359

## RESULT 12

US-09-844-705-13  
; Sequence 13, Application US/09844705  
; Patent No. US20020133836A1  
; GENERAL INFORMATION:  
; APPLICANT: BOSWORTH, BRAD  
; APPLICANT: VOGELI, PETER  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY  
; TITLE OF INVENTION: RESISTANT TO F18 E. COLI ASSOCIATED DISEASES  
; FILE REFERENCE: 21419/90368  
; CURRENT APPLICATION NUMBER: US/09/844,705  
; CURRENT FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: 09/443,766  
; PRIOR FILING DATE: 1999-11-19  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 365  
; TYPE: PRT  
; ORGANISM: Porcine  
US-09-844-705-13

Query Match 55.1%; Score 1044.5; DB 10; Length 365;  
Best Local Similarity 62.6%; Pred. No. 8e-100;  
Matches 194; Conservative 38; Mismatches 75; Indels 3; Gaps 2;

Qy 17 PMTQMSSGNTESPEMRDSEHGNGELRGMTFTNSIGRLGNQMGGEYATLALARMNGRL 76  
Db 51 PVALIFCLAGTPVHPNADSCPKH-PASFGTWTIYDPGRFGNQMGQYATLLALAGLNGRQ 109  
Qy 77 AFIPASMNALAPIRISLPVLHSDTAKKIPQNYHLNDMMEERYRHPGHVFRFTGYPC 136  
Db 110 AFIQPAMHVLAPVFRITLPVLAPEVDHAPWRELELDHWMSEDYAHLKEPWLKLTGPPC 169  
Qy 137 SWTFVYHLRPEILKEFTLHDHVRREAQAFRLGLRV--NGSOPSTFVGHVVRGDDYVHVMP 194  
Db 170 SWTFVYHLRPEILKEFTLHDHVRREAQAFRLGLRV--NGSOPSTFVGHVVRGDDYVHVMP 229  
Qy 195 NWKGVVADRGYLEKALDMFRARYSSPVFTVTSNGMACRENINASRGDVFVAGNGIEGS 254  
Db 230 KRWKGVVGDGYLQAMDWFARYEAPVFTVTSNGMEWCKENIDTSRGDVFVAGNGIEGS 289  
Qy 255 PAKDFALLTQCNHMTIGTGFVWAAYLAGGDTIYLANITLPSPLKVKFKEAAFLPEW 314  
Db 290 PARDFALLVQCNHMTIGTGFVWAAYLAGGDTIYLANITLPSPLKVKFKEAAFLPEW 349  
Qy 315 VGIPADLSPL 324  
Db 350 VGINADLSPL 359

Wed May 28 09:21:53 2003

Db 51 PVAIFCLAGIPVHPNADSDCPKH-PASFSGTWITYPDGRFCNQMGQYATLALAQINRG 109  
 Qy 77 AFIPASHMNAIAPRISLPLVHSDTAKKIPWQNYLNDWMEERYRHPGHVFTGYPC 136  
 Db 110 AQIOPAMHVAIAPVFTITPLVLAPEVDHAPWRELEHDMSEDAHLKEPWLKLTGPPC 169  
 Qy 137 SWTFVHHLRPEILKEFTLHDHVRREAAFLRGLV--NGSOPSTFVGVHVRGDIYVHVP 194  
 Db 170 SWTFPHHLRPEILKEFTLHDHVRREAAFLRGLV--NGSOPSTFVGVHVRGDIYVHVP 229  
 Qy 195 NVWKGVDADRGYLEKALDMFRARYSSPVFVTSNGMWCNENINASRGDVVFAGNGIEGS 254  
 Db 230 KWKGVGVDGRYLOQADWFRARYEAPVFTVTSNGMWCNENINASRGDVVFAGNGIEGS 289  
 Qy 255 PAKDRALLTOCHNTMTTGTGFWAAYLAGGDIYLYANTLPDPSFLVKVFKPEAAFLPEW 314  
 Db 290 PAKDRALLTOCHNTMTTGTGFWAAYLAGGDIYLYANTLPDPSFLVKVFKPEAAFLPEW 349  
 Qy 315 VGIPADLSPL 324  
 Db 350 VGINADLSPL 359  
 RESULT 13  
 US-09-848-838-2  
 ; Sequence 2, Application US/09648836  
 ; Patent No. US20020037570A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Taylor, Diane  
 ; APPLICANT: Wang, Ge  
 ; APPLICANT: Palcic, Monica  
 ; TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE  
 ; FILE REFERENCE: 07254-061002  
 ; CURRENT APPLICATION NUMBER: US/09/848,838  
 ; PRIOR FILING DATE: 2001-05-03  
 ; PRIOR FILING DATE: 1999-11-02  
 ; PRIOR FILING DATE: 1998-11-04  
 ; NUMBER OF SEQ ID NOS: 23  
 ; SOFTWARE: FastSeq for Windows version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 300  
 ; TYPE: PRT  
 ; ORGANISM: Helicobacter pylori  
 US-09-848-838-2

Query Match 7.48; Score 141; DB 10; Length 300;  
 Best Local Similarity 24.0%; Pred. No. 2.9e-06;  
 Matches 81; Conservative 33; Mismatches 125; Indels 98; Gaps 14;  
 Qy 54 GRGNMGGEYA-----TLFALARMNGRLAFIPASMHNAIAPRIS-- 94  
 Db 10 GGLGNMGQYAFKSLQXHLNTPVLLDTSTFDSNKKWQLEFPIDLPYANAKETAIAM 69  
 Qy 95 --LPVLHSDTAKKIPWQNYLNDWMEERYRHI-PGHVFTGYPCSWTFVHHLRPEILKE 151  
 Db 70 QHLPKLVDRALKYIGFDVRSQIVPEYEPKLPKSLRTYFFGYFODPRYFDAISSLIKOT 129  
 Qy 152 FTL-----HDHVRREAAFLRGLRVNGSOPSTFVGVHVRGDIYVHVPNVKGVYA 202  
 Db 130 FTLPPLPPNNKNNKEEYQKRLS--LILAKNSVF--VHVRGDIYVGI-----GCOL 179  
 Qy 203 DRGYLEKALDMFRARYSSPVFVTSNGMWCNENINASRGDVVFAGNGIEGSPAKDFA-- 260  
 Db 180 GIDYQKALEYMAKRVNMFELV-----FCE-----DLKFTQNLDLGYPTDMITR 225  
 Qy 261 -----LITQCNHTMTTGTGFWAAYLAGGDIYLYANTLPD-----SPFLKVEK 305  
 Db 226 DKEEAYWMLMCSCKHGIANSYSTYSWAAAYL-----MEN---PEKIIIGKHWLFG 275  
 Qy 306 PEAAFLPEWVGIPADLSPLKALTPACPRSHHLKAK 342

Db 276 HENILCKREWVKI-----ESHFEVKSQ 296  
 RESULT 14  
 US-10-214-524-25  
 ; Sequence 25, Application US/10214524  
 ; Publication No. US20030073142A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chen, Swey-Shen Alex  
 ; APPLICANT: Yang, Yong-Min  
 ; APPLICANT: Barankiewicz, Theresa J.  
 ; APPLICANT: Chen, Zhong  
 ; TITLE OF INVENTION: IMMUNOGLOBULIN E VACCINES AND METHODS OF USE THEREOF  
 ; FILE REFERENCE: IGE-00101.P.1.1  
 ; CURRENT APPLICATION NUMBER: US/10/214,524  
 ; CURRENT FILING DATE: 2002-08-08  
 ; PRIOR FILING DATE: 2001-08-13  
 ; NUMBER OF SEQ ID NOS: 61  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 25  
 ; LENGTH: 496  
 ; TYPE: PRT  
 ; ORGANISM: Cat (Felis catus)  
 US-10-214-524-25  
 Query Match 5.2%; Score 99; DB 9; Length 496;  
 Best Local Similarity 25.0%; Pred. No. 0.14;  
 Matches 55; Conservative 21; Mismatches 84; Indels 60; Gaps 12;  
 Qy 126 GHFVRFTGYPCSWTFVHHLRPEILKEFTLHDHVRREAAFLRGLRVNGSOPSTFVGVHVR 185  
 Db 249 GEWVSQKTYTCQVY-----OGTFEDHARKCTESDPRGVSILSPSPSL----- 293  
 Qy 186 RGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGM--AWCENINASRGD 243  
 Db 294 -DLYVHSPKI-ICLVVD-----LANTDGMILTWSREN----- 324  
 Qy 244 VVFAGNGIEGSPAKDFAITQCNHTMTTGTGFWAAYLAGGDIYLYANTLPDSP--F 300  
 Db 325 ---GESVHPDP---MVKTKYNGTITVTSTPLPVDATDWDVEGT-YOCKVTHPDLPKDIV 376  
 Qy 301 LKVPK-PEAAFLPE-WVGIPADLSPLK-ALTPACPRSHF 337  
 Db 377 RSIAPKGRPEPEVYVFLPPEGEPEKTKDKVTLTCLQNF 416

RESULT 15  
 US-09-479-614-14  
 ; Sequence 14, Application US/09479614  
 ; Publication No. US20030013183A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McCall, Catherine  
 ; APPLICANT: Weber, Eric  
 ; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods  
 ; FILE REFERENCE: P-1047  
 ; CURRENT APPLICATION NUMBER: US/09/479,614  
 ; CURRENT FILING DATE: 2000-01-07  
 ; EARLIER APPLICATION NUMBER: 60/115,033  
 ; EARLIER FILING DATE: 1999-01-07  
 ; NUMBER OF SEQ ID NOS: 34  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 14  
 ; LENGTH: 431  
 ; TYPE: PRT  
 ; ORGANISM: Felis catus  
 US-09-479-614-14  
 Query Match 5.1%; Score 97.5; DB 9; Length 431;  
 Best Local Similarity 25.0%; Pred. No. 0.16;  
 Matches 52; Conservative 19; Mismatches 78; Indels 59; Gaps 11;  
 Qy 126 GHFVRFTGYPCSWTFVHHLRPEILKEFTLHDHVRREAAFLRGLRVNGSOPSTFVGVHVR 185

Db	184	GEWYSQTYTCQVITY	: :       : :       : :       : :	228
QY	186	RGDYVHVMPNWKGVVADRGYLEKALDMPARVSSPVFVTSNGM--ACNRENINASRGD	---QQFTEDHARKCTESDPRGVSYLSPPSP---	243
Db	229	-DLVHKSPKI-TCLVVD-	: :       : :       : :	259
QY	244	VVFAGNIEGSPAKDFALLTQCNIHTMTCTFGIWAAYLAGDITYLANVTLPDSP--F	---LANTDGMILTWSEN---	300
Db	260	-----GESVHPDP---MVKKQYNGTIVTSTLFDVADTWVEGET-YQCKVTHPDLPKQIV	: :       : :       : :	311
QY	301	LKVFK-PEAAFLPE-VWGPADLSPLLK	326	
Db	312	RSTAKAPGRFPPEVYVFLPPEGEPTK	: :       : :	339

Search completed: May 27, 2003, 15:25:43  
Job time : 16.2516 secs

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:00:45 ; Search time 12.7916 Seconds

(without alignments)

2652.940 Million cell updates/sec

Title: US-10-040-863-10

Perfect score: 1896

Sequence: 1 LQQRIVKLOPLSEKELPMTT.....RSHFHLKAKGVTCYVAGRAF 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

.Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1386.5	73.1	343	2 A56098	alpha(1,2)fucosylit
2	1291	68.1	354	2 B56392	beta-galactoside a
3	1080.5	57.0	373	2 A56392	beta-galactoside a
4	1051.5	55.5	365	2 A36047	galactoside 2-alph
5	846	44.6	159	2 S46494	galactoside 2-alph
6	581	30.6	142	2 S51582	galactoside 2-alph
7	214.5	11.3	281	2 T44328	hypothetical prote
8	183	9.7	443	2 H87911	protein B0205.4 [i
9	170.5	9.0	363	2 T20745	hypothetical prote
10	154	8.1	348	2 T31916	hypothetical prote
11	153.5	8.1	335	2 T32294	hypothetical prote
12	139.5	7.4	383	2 T20572	hypothetical prote
13	138.5	7.3	500	2 T22068	hypothetical prote
14	133.5	7.0	299	2 H71976	hypothetical prote
15	131	6.9	365	2 T25309	hypothetical prote
16	127	6.7	388	2 T32307	hypothetical prote
17	127	6.7	392	2 T25334	hypothetical prote
18	124	6.5	353	2 T25390	hypothetical prote
19	120	6.3	625	2 T21051	hypothetical prote
20	119	6.3	434	2 T26275	hypothetical prote
21	118.5	6.2	381	2 T15140	hypothetical prote
22	117.5	6.2	365	2 T32253	hypothetical prote
23	111	5.9	395	2 T32309	hypothetical prote
24	109.5	5.8	371	2 T32692	hypothetical prote
25	109	5.7	277	2 T25307	hypothetical prote
26	98.5	5.2	317	2 T33887	hypothetical prote
27	96.5	5.1	355	2 T34405	hypothetical prote
28	93.5	4.9	348	2 T02798	hypothetical prote
29	93	4.9	597	2 G87258	acyl-CoA dehydroge

ALIGNMENTS

RESULT 1

A56098

alpha(1,2)fucosyltransferase Sec2, long form - human

C:Species: Homo sapiens (man)

C:Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 21-Jul-2000

C:Accession: A56098

R:Keilly, R.J.; Rouquier, S.; Giorgi, D.; Lennon, G.G.; Lowe, J.B.

J. Biol. Chem. 270, 4640-4649, 1995

A:Title: Sequence and expression of a candidate for the human Secretor blood group al

tes with the non-Secretor phenotype.

A:Reference number: A56098; MUID:95181460; PMID:7876235

A:Accession: A56098

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-343 <REL>

A:Cross-references: GB:U17894; NID:9687618; PIDN:AAC24453.1; PID:9687619

C:Genetics:

A:Gene: GDB:PUT2; SE

A:Cross-references: GDB:120619; OMIM:182100

A:Map position: 19q13.3-19q13.3

C:Keywords: glycoprotein; Golgi apparatus; transmembrane protein

Query Match

Best Local Similarity 79.1%; Score 1386.5; DB 2; Length 343;

Matches 258; Conservative 26; Mismatches 31; Indels 11; Gaps 2;

QY 1 LQQRIVKLOPLSEKELPMTTOMSSNGTESPEMRDSEQHNGELRGFTINISIGRLGNQM 60

DB 28 VOORLAKIQAM--WELPV-----QIPVLASTSKALGPSOLRGMTWINAIGRLGNQM 76

QY 61 GEVATLFAIARMNGRLAFIPASMHNNAIATFRISLPVLSHSDTAKKIPWNYHLNDKKEER 120

DB 77 GEYALYALAKMNGRPAFPAQMHSTLAPITFLPVLHSATASRIPWQNYHLNDKKEE 136

QY 121 YRHIPGHFVRYTGYPCSWTFYHHLRPEILKEFTLHDHVEEACAFRLGRVNGSQSTFV 180

DB 137 YRHIPGEYVRYTGYPCSWTFYHHLRQEILOEFTLHDHVEEACAFRLGRVNGSQSTFV 196

QY 181 GVHVARGDYVHVYMPNVKGVADRGYLEKALDMFRARYSSPVVVTNSGMWCRENINAS 240

DB 197 GVHVARGDYVHVYMPNVKGVADRGYLEKALDMFRARYSSPVVVTNSGMWCRENIN 256

QY 241 RGDVVFAGNGIEGSPAKDFALLTCQNHMTTGTGICRAYLAGDGTIYLANITLPDSPF 300

DB 257 RGDVVFAGNGIEGSPAKDFALLTCQNHMTTGTGICRAYLAGDGTIYLANITLPDSPF 316

QY 301 LKVEKPEAAFLPEWVGIPADLSPLLK 326

DB 317 LKIFKPEAAFLPEWVGIPADLSPLLK 342

RESULT 2



```
B56392
beta-galactoside alphas,2-fucosyltransferase II - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 05-Nov-1999
C:Accession: B56392
R:Hitoshi, S.; Kusunoki, S.; Kanaazawa, I.; Tsuji, S.
J. Biol. Chem. 270, 8844-8850, 1995
A:Title: Molecular cloning and expression of two types of rabbit beta-galactoside alphas
A:Reference number: A56392; MUID:95238380; PMID:7721792
A:Accession: B56392
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-354 <HIT>
A:Cross-references: GB:X80225; NID:g854356; PIDN:CAA56512.1; PID:g854357
C:Keywords: transmembrane protein

Query Match 68.1%; Score 1291; DB 2; Length 354;
Best Local Similarity 83.9%; Pred. No. 1.4e-100; Indels 0; Gaps 0;
Matches 235; Conservative 20; Mismatches 25;

QY 46 GMETINSIGRLNOMGEYATLFAALARMNGRLAFIPASMHNALAPIFRISLPVLSHTAKK 105
DB 71 GWTINAMGLNOMGEYATLALAKENGRPAIIPAQMHSTLAPIFRISLPVLSHTASR 130
QY 106 IPQNYHLNDWMEERYRHIPGHVFTGPCSWTFYHHLRPEILKEFTLHDHVRERAAQAF 165
DB 131 VPQNYHLNDWMEERYRHIPGVFVRLTGPCSWTFYHHLRPEILKEFTLHDHVRERAAQAF 190
QY 166 LRLRVNGSQPSTFVGVHVRGDDYVHVMNVKGVVADRGYLEKALDMFRARYSSPVFV 225
DB 191 LRLRVNGSRPSTFVGVHVRGDDYVHVMNVKGVVADRGYLEKALDMFRARYSSPVFV 250
QY 226 TSNMGAWCRENINASRGDVVAGNGLEGSPAKDFALLTCQNHPTMTIGTFGIWAAYLAGG 285
DB 251 TSNMGAWCRENIDASRGDVVAGNGLEGSPAKDFALLTCQNHPTMTIGTFGIWAAYLAGG 310
QY 286 DTIYLANITLPDSFPLKVKPEAAFLPEWVGIPADLSPLL 325
DB 311 DTIYLANITAPDSFPLKVKPEAAFLPEWVGITANMGAL 350

RESULT 3
A56392
beta-galactoside alphas,2-fucosyltransferase I - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 05-Nov-1999
C:Accession: A56392
R:Hitoshi, S.; Kusunoki, S.; Kanaazawa, I.; Tsuji, S.
J. Biol. Chem. 270, 8844-8850, 1995
A:Title: Molecular cloning and expression of two types of rabbit beta-galactoside alphas
A:Reference number: A56392; MUID:95238380; PMID:7721792
A:Accession: A56392
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-373 <HIT>
A:Cross-references: GB:X80226; NID:g854354; PIDN:CAA56513.1; PID:g854355
C:Keywords: transmembrane protein

Query Match 57.0%; Score 1080.5; DB 2; Length 373;
Best Local Similarity 63.3%; Pred. No. 6.9e-83;
Matches 200; Conservative 44; Mismatches 51; Indels 21; Gaps 3;

QY 16 LPMTTOMS-----SQNTESPENRRDSEHGNGELRGMTINSIGRLNOMGEYATLALA 70
DB 58 LPTVSPASNASSCAGRPAAPS-----GIWTHPDGRFGNOMGGYATLLALA 103
QY 71 RNGLAFTPASMHNALAPIFRISLPVLSHTAKKIPQNYHLNDWMEERYRHIPGHFVR 130
DB 104 QLNGRFAFLPAMHAALAPVFRITLPVLAPEVNNRTSKQLLLHOMSEYSLDPTLK 163
QY 131 FTGYPGCSWTFYHHLRPEILKEFTLHDHVRERAAQAFRLGRV--NGSQPSTFVGVHVRGDD 188
DB 164 FTGFPSCSWTFYHHLRPEILKEFTLHDHVRERAAQAFRLGRV--NGSQPSTFVGVHVRGDD 223

189 YVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTNSMGAWCRENINASRGDVV 248
224 YLQVMPQPNKGVVGDRAVLAQMDWFRARHEAPFVVTNSMGKWCWENIDASRGDWFAG 283
249 NGIEGSPAKDFALLTCQNHPTMTIGTFGIWAAYLAGGDTIYLANITLPDSFPLKVKPEA 308
284 NGLSSPAKDFALLTCQNHPTMTIGTFGIWAAYLAGGDTIYLANITLPDSFPLKVKPEA 343
309 AFLPEWVGIPADLSPL 324
344 AFLPEWVGINADLSPL 359

RESULT 4
A36047
galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) - human
C:Species: Homo sapiens (man)
C>Date: 16-Nov-1990 #sequence_revision 16-Nov-1990 #text_change 10-Sep-1997
C:Accession: A36047
R:Larsen, R.D.; Ernst, L.K.; Nair, R.P.; Lowe, J.B.
Proc. Natl. Acad. Sci. U.S.A. 87, 6674-6678, 1990
A:Title: Molecular cloning, sequence, and expression of a human GDP-L-fucose:beta-D-g
A:Reference number: A36047; MUID:90370848; PMID:2118655
A:Accession: A36047
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-365 <LAR>
A:Cross-references: GB:M35531; NID:g183887; PID:g306930
C:Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein

Query Match 55.5%; Score 1051.5; DB 2; Length 365;
Best Local Similarity 63.3%; Pred. No. 1.8e-80;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

QY 25 GNTSEPMRRDSEHGNGELRGMTINSIGRLNOMGEYATLALAGLNRRAFILPAMH 84
DB 59 GTAMGPNASSCCPCH-PASLSGTWTVYVNGRFGNOMGGYATLLALAGLNRRAFILPAMH 117
QY 85 NALAPIFRISLPVLSHTAKKIPQNYHLNDWMEERYRHIPGHVFRFTGYPGCSWTFYHHL 144
DB 118 AALAPVFRITLPVLAPEVDSRTPRRELQHLNDWMEERYADRLDPLKLSGFPSCSWTFPHHL 177
QY 145 RPEILKEFTLHDHVRERAAQAFRLGRV--NGSQPSTFVGVHVRGDDYVHVMNPVKGVA 202
DB 178 RQIRERFTLHDHVRERAAQAFRLGRV--NGSQPSTFVGVHVRGDDYVHVMNPVKGVA 237
QY 203 DRGYLEKALDMFRARYSSPVFVVTNSMGAWCRENINASRGDVVAGNGLEGSPAKDFALL 262
DB 238 DSAYLRQMDWFRARHEAPFVVTNSMGAWCRENINASRGDVVAGNGLEGSPAKDFALL 297
QY 263 TQCNHTMTIGTFGIWAAYLAGGDTIYLANITLPDSFPLKVKPEAAFLPEWVGIPADLS 322
DB 298 TQCNHTMTIGTFGIWAAYLAGGDTIYLANITLPDSFPLKVKPEAAFLPEWVGIPADLS 357
QY 323 PLLKALTP 330
DB 358 PLWTLAKP 365

RESULT 5
S46494
galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) B - rat (fragment)
N:Alternate names: alpha-1,2-fucosyltransferase
C:Species: Rattus norvegicus (Norway rat)
C>Date: 26-Dec-1994 #sequence_revision 23-Feb-1996 #text_change 08-Oct-1999
C:Accession: S46494
R:Piau, J.P.; Labarriere, N.; Dabouis, G.; Denis, M.G.
Biochem. J. 300, 623-626, 1994
A:Title: Evidence for two distinct alpha-(1,2)-fucosyltransferase genes differentiall
A:Reference number: S46493; MUID:94280382; PMID:8010942
A:Accession: S46494
A:Molecule type: mRNA
```



Db 217 DPVRKRVSEWNIYWIHPTNHRKKPEKSTVSIFSVTFQLKRVDFLENVGSLTVRNAR 276  
 QY 171 V-----NGSQ-----PSTFVGHHVVRGDYVHV--MPNVWKGVADRGYLEKAL 211  
 Db 277 VIETNVANDQALELPEEDAFKTNMVGWHIRHGMDISMSNSRNLHGHDVDPTEIYKRAI 336  
 QY 212 DMFRARYSPVFTVTSNGMAWCRENINASRGDVFVAGNGIEGSPAKDFALLTQCNTHTMT 271  
 Db 337 QOISKIYENVAFIQSDNVAVARRNLKLGKETLHF---CPGPREVDMAILKSCDSVIIS 393  
 QY 272 IGTGCIWAAYL---AGGDTIYLANVTLTDSPLKVKFKPEAAFLPEWGI 317  
 Db 394 TGTGQWWSAYLNVNASPDVYVYKHPAGVSMVEKMTNKTFLPKSWTAL 442

RESULT 9  
 T20745  
 hypothetical protein F11A5.5 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000  
 C:Accession: T20745  
 R:Gardner, A. submitted to the EMBL Data Library, March 1997  
 A:Reference number: Z19319  
 A:Accession: T20745  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-363 <W1>  
 A:Cross-references: EMBL:292830; PIDN:CAB07352.1; GSPDB:GN00023; CESP:F11A5.5  
 A:Experimental source: clone F11A5  
 C:Genetics:  
 A:Gene: CESP:F11A5.5  
 A:Map position: 5  
 A:Introns: 42/1; 55/2; 136/3; 175/3; 210/2; 243/1; 299/1; 330/3  
 C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 9.0%; Score 170.5; DB 2; Length 363;  
 Best Local Similarity 23.4%; Pred. No. 1.4e-06;  
 Matches 73; Conservative 45; Mismatches 101; Indels 93; Gaps 17;

QY 52 SIGRLNGMGEYATLAFALARMNGRLA--RIPASMH-----NALAPIFRISLPVLHSDT 102  
 Db 78 SSSRLGNHLFELASVLSISRELQVRPTFTIENCYHEKMWEDSNLPLGLMNHFLINGSV 137  
 QY 103 AKKIPQWYH-----LNDWMEERYRHIPGHVFTGYPCSWTFYHHLRPEILKEF 152  
 Db 138 PSSVKRVKHEKQCTFDDPSLLDNYDEYHLTG-----THYQ-SWKYFSHMRNELI--- 188  
 QY 153 TLHDHVREEAQAFRLGLRVNGSQPSTFVG-VHVRGDYVHVPNVWKGV-VADRGYLEKA 210  
 Db 189 ---GYLKTENTY---MDLPKSGENTFITCVHVRRGDFLRV-----GFHVADENTIRSS 236  
 QY 211 LDMFRARYSPVFTVTSNGMAWCRENINASRGDVFVAGNGIE----- 252  
 Db 237 LNLISQVAKRANTAT-----VFGDDYEFMDSLNRKTSINAFVSQ 278  
 QY 253 GSPAKDFALL-TQCNHTMTI--GTGCIWAAYLAGGDTIYLANVTLTDSPLK----- 302  
 Db 279 NSPADLLYAKSCDVVLIITAAHSTFGWGWGYFSKGNRYV---YT---DIQFTKDWILETG 333  
 QY 303 VKFPEAAFLPEW 314  
 Db 334 EFISEDYILPHW 345

RESULT 10  
 T31916  
 hypothetical protein C17A2.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jul-2000  
 C:Accession: T31916  
 R:Sammons, L.; Wohldmann, P.  
 submitted to the EMBL Data Library, July 1997

A:Description: The sequence of C. elegans cosmid C17A2.  
 A:Reference number: Z21098  
 A:Accession: T31916  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-348 <SAM>  
 A:Cross-references: EMBL:AF016654; PIDN:AAB66030.1; GSPDB:GN00020; CESP:C17A2.4  
 A:Experimental source: strain Bristol N2; clone C17A2  
 C:Genetics:  
 A:Gene: CESP:C17A2.4  
 A:Map position: 2  
 A:Introns: 94/3; 133/3; 168/2; 272/1  
 C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 8.1%; Score 154; DB 2; Length 348;  
 Best Local Similarity 22.8%; Pred. No. 3.2e-05;  
 Matches 63; Conservative 53; Mismatches 102; Indels 58; Gaps 15;

QY 52 SIGRLNGMGEYATLAFALARMNGR--LAFIPASMHNALAPIFRISLP-----VLHSDT 102  
 Db 36 SPSRLGNHIFEFASLFLSERLHRTPLFLVNEFHOKMLDETARKVMPGLVEKFTVINGSL 95  
 QY 103 AKKI---PHQ-----NYHLNDWMEERYRHIPGHVFTGYPCSWTFYHHLRPEILKEF 152  
 Db 96 PKSIKOTPFQKVCCHRENPEILEKIDKYLHTGMFYQ-----SWKYPNMQEQLL--- 146  
 QY 153 TLHDHVREEAQAFRLGLRVNGSQPSTFVG-VHVRGDYVHVPNVWKGV-VADRGYLEKAL 211  
 Db 147 ---DFLDDSSQDF--GNLPSRNOQTHVTCVHARRGFDVY-----GFOAADPDFIRNSV 195  
 QY 212 DMFRARYSPV-----FVTSNGMAWC---EN--INASRGDVFVAGNGI--EGSPA 256  
 Db 196 KYAENFIPEIEYKVKVRKVIFGDDLEFMSLFENVSVSTDEPEYMFPAEYISQNSPA 255  
 QY 257 KQFALLTQ--CNHTMTI--GTGCIWAAYLAGGDTIY 289  
 Db 256 EDLTSKQNCDIVLISAPKSTFGWGWGYFSKGNKV 291

RESULT 11  
 T32294  
 hypothetical protein K06H6.6 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 20-Jun-2000  
 C:Accession: T32294  
 R:David, M.; Wohldmann, P.; Bauer, C.; Clarke, K.  
 submitted to the EMBL Data Library, September 1997  
 A:Description: The sequence of C. elegans cosmid K06H6.  
 A:Reference number: Z21147  
 A:Accession: T32294  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-335 <DAV>  
 A:Cross-references: EMBL:AF024500; PIDN:AAB70365.1; GSPDB:GN00023; CESP:K06H6.6  
 A:Experimental source: strain Bristol N2; clone K06H6  
 C:Genetics:  
 A:Gene: CESP:K06H6.6  
 A:Map position: 5  
 A:Introns: 17/1; 32/2; 58/1; 111/3; 150/3; 184/2; 218/1; 243/3; 272/1; 303/3  
 C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 8.1%; Score 153.5; DB 2; Length 335;  
 Best Local Similarity 22.8%; Pred. No. 3.3e-05;  
 Matches 77; Conservative 50; Mismatches 118; Indels 93; Gaps 19;

QY 24 SGNTESPEMRDSEHQNGELRGMTINSIGRLNGMGEYATLAFALARMNGRLA--PIPA 81  
 Db 27 SATAEIPPLRQESFLMKKRLSSRMA--PTARLGNHMFELAAVLGIGSRMLNRTATTFIED 84  
 QY 82 SMHNALAPIFRISLP-----VLHSDTAKKIPW--QNYHLN-----DWMEER 120  
 Db 85 EIYRRMTESKRAIPGLVGGFEILING---KVPLYIKTKLNTKRCVDFVPLTHEHNDE 140



```
QY 54 GRLGNMGGEYA-----TLFALRMNGRLAFIPASMHNALAPIERIS-- 94
||| ||| |||
Db 10 GGLGNMGFOYAFKSLQKHSNTPVLIDITSDGSKNRKQLELPDIDLPSAKETAIAKM 69
||| ||| |||
QY 95 --LPVLHSDTAKKIPQNNVHLNDWMEERYHI-PGHEVRFYTCYPCSWTFYHHLRPEILKE 151
||| ||| |||
Db 70 QHLPKLVDRDALKYMGRDVSQEIIVEYEPKLLKSLRTYFYGFQDPYFQDAISSLIKQT 129
||| ||| |||
QY 152 FTL-----HDHVREAAQAFRLGLR-VNSQSPSTFVGTVHVRGRDYYVHVMNVKGVVADRG 205
||| ||| |||
Db 130 FTLPPPPNGNNKKKEEYHKKLSLILAAKNVF--AHLRRGDYVGI-----GCOLGID 181
||| ||| |||
QY 206 YLEKALDMFRARYSSPVFVVTSGMAWCRENINASKRGDVFAGWGIEGSPAKDFA----- 260
||| ||| |||
Db 182 YOKKAVEYMAKVPNMELFV-----FCE-----DLKFTQNLDLGYFPFMDMTTRDKD 227
||| ||| |||
QY 261 -----LLTQCNTHITMTIGTGWAAAYLAG--GDTIYLANYTLPDSPFLKVKFPEAAF 310
||| ||| |||
Db 228 EEAYDMLLMQSKHGIIANSTYSWAAVLIINPGKIIT-----GPKHMLFGHENIL 279
||| ||| |||
QY 311 LPEWVGIPADLPLKALTTPACPRSHFHLKAK 342
||| ||| |||
Db 260 KEWKI-----ESHFEVKSQ 295
||| ||| |||

RESULT 15
T25309
hypothetical protein T26E4.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C:Accession: T25309
R:McMurray, A.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z20013
A:Accession: T25309
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-365 <WIL>
A:Cross-references: EMBL:Z81132; PIDN:CAB03434.1; GSPDB:GNO00023; CESP:T26E4.4
A:Experimental source: clone T26E4
C:Genetics:
A:Gene: CESP:T26E4.4
A:Map position: 5
A:Introns: 50/1; 80/2; 140/3; 179/3; 218/2; 253/1; 278/1; 306/3; 325/1; 342/3
C:Supertfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 6.9%; Score 131; DB 2; Length 365;
Best Local Similarity 20.7%; Pred. No. 0.0028;
Matches 61; Conservative 43; Mismatches 106; Indels 84; Gaps 13;

QY 57 GNMGEYATLFLARMNGRLAFIPR-----SMENALAPIERISLPVLHSDTAKKIPWQ 109
||| ||| |||
Db 86 GNKLEIISLLGIANSLQRRPVIDATIFSNTSLKSKTQPLEP---KLVEQFDLKNIPAS 142
||| ||| |||
QY 110 NY--HLNDWME-----ERYKHIPGHVRFRTGYPCSWTFYHHLRPEILKEFT 153
||| ||| |||
Db 143 SVSSHQMNWVKCCIEDDPKPKMLNRSEQHMLMNGHYFQ-----SPKYFHHLRSEIREWLA 196
||| ||| |||
QY 154 LHDHVREAAQAFRLGLRVNSQSPSTFVGTVHVRGRDY----VHVMNVKGVVADRGYLEX 209
||| ||| |||
Db 197 PSKMAKLAETVL-----TSELKEDIICTHIRRGDFQTDGVH-QP-----SDPNFTRA 244
||| ||| |||
QY 210 ALDMFRARYSS---PVFVVTSGMAWCRENI-----NASR 241
||| ||| |||
Db 245 ATDFLVKKYQKWHYRITVVVFQNDVNFSKAVFEDRVNSSVIPNRTTTPNPIPENSXP 304
||| ||| |||
QY 242 GDWTFAGNIEGSPAKFALLTQCNTIMTICTFGIWAAYLAGGDTIYLIANYTL 295
||| ||| |||
Db 305 YSVILPQN---STPENDLAFSRQAPSS-----TFGWLSYIAKRSAYVLRPFCL 350
||| ||| |||
```

Search completed: May 27, 2003, 15:08:18  
Job time : 13.7916 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:53:11 ; Search time 10.3317 Seconds

(without alignments)

1417.108 Million cell updates/sec

Title: US-10-040-863-10

Perfect score: 1896

Sequence: 1 LQQRIVKLPFLSEKELPMNT.....RSHFLKARKGVTCYVAGRAF 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1430	75.4	340	1	FUT2_PIG
2	1393	73.5	344	1	FUT2_BOVIN
3	1386.5	73.1	343	1	FUT2_HUMAN
4	1338	70.6	368	1	FUT2_MOUSE
5	1291	68.1	354	1	FUT2_RABIT
6	1105.5	58.3	376	1	FUT1_MOUSE
7	1101.5	58.1	376	1	FUT1_RAT
8	1080.5	57.0	373	1	FUT1_RABIT
9	1051.5	55.5	365	1	FUT1_HUMAN
10	1050.5	55.4	365	1	FUT1_PIG
11	846	44.6	159	1	FUT2_RAT
12	168	8.9	365	1	YKQ7_CAEEL
13	91	4.8	442	1	VGLD_HSVBK
14	89	4.7	353	1	FNAB_PORGI
15	89	4.7	773	1	CDH1BACH
16	89	4.7	2058	1	MT10_HUMAN
17	88.5	4.7	359	1	PN7_RAT
18	88.5	4.7	402	1	VGLD_HSVBA
19	88.5	4.7	452	1	VGLD_HSVBA
20	88	4.6	217	1	GRB2_MOUSE
21	87	4.6	217	1	GRB2_HUMAN
22	87	4.6	438	1	NOOL_THERM
23	87	4.6	533	1	QUTD_EHENT
24	86.5	4.6	1112	1	CN3B_HUMAN
25	85.5	4.5	738	1	PL03_HUMAN
26	85.5	4.5	1928	1	LPH_RAT
27	84.5	4.5	344	1	RIBB_HELPY
28	84.5	4.5	383	1	AAT_THEAD
29	84.5	4.5	718	1	ATPC_MYCTU
30	84	4.4	1780	1	POIG_MVEV
31	83.5	4.4	273	1	TC1A_CAEEL
32	83.5	4.4	350	1	FMB2_FORGI
33	83.5	4.4	440	1	CYE_PARDE

## RESULT 1

FUT2\_PIG STANDARD; PRT; 340 AA.

AC Q10982: Q29044; O19100;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)F2)  
DE D-galactoside 2-alpha-L-fucosyltransferase 2)  
GN FUT2.

OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97468270; PubMed=9321466;  
RA Meljerkink E., Fries R., Voegeli P., Masabanda J., Wigger G., Stricker C., Neuenschwander S., Bertschinger H.U., Stranzinger G.;  
RT "Two alpha(1,2) fucosyltransferase genes on porcine chromosome 6c11 are closely linked to the blood group inhibitor (S) and Escherichia coli F18 receptor (ECF18R) loci.";  
RT Mamm. Genome 8:736-741(1997).  
RL [2]  
RP SEQUENCE FROM N.A.  
RX Conney S., Mountouris E., McKenzie I.F.C., Sandrin M.S.;  
RT "Molecular cloning and characterization of the pig secretor type alpha(1,2)fucosyltransferase.";  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 69-334 FROM N.A.  
RA Petit J.M.;  
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE OF 69-80; 119-133 AND 316-334.  
RC TISSUE=Submaxillary gland;  
RX MEDLINE=96064742; PubMed=7592879;  
RA Thurnin J., Blaszyk-Thurin M.;  
RT "Porcine submaxillary gland GDP-L-fucose: beta-D-galactoside alpha-2-L-fucosyltransferase is likely a counterpart of the human secretor gene-encoded blood group transferase.";  
RL J. Biol. Chem. 270:26577-26580(1995).  
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP + alpha-L-fucosyl-1,2-beta-D-galactosyl-R.  
CC -!- PATHWAY: Glycosylation.  
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

O61409 mus musculus  
P76464 escherichia  
P45784 vibrio chol  
P72933 synechocyst  
P87503 human adeno  
P7846 rattus norv  
P78357 homo sapien  
P87379 xenopus lae  
Q82140 helicobacte  
P27339 rattus norv  
P47527 mycoplasma  
O54991 mus musculu

34 83.5 4.4 799 1 CN3B\_MOUSE  
35 83.5 4.4 1534 1 YFAS\_ECOLI  
36 83 4.4 252 1 GSPN\_VIECH  
37 82.5 4.4 257 1 GLO2\_SYNY3  
38 82.5 4.4 1193 1 DPOL\_ADE04  
39 82.5 4.4 1381 1 CTAL\_RAT  
40 82.5 4.4 1384 1 CTAL\_HUMAN  
41 82 4.3 217 1 GRB2\_XENLA  
42 82 4.3 344 1 RIBB\_HELPY  
43 82 4.3 1840 1 SUIS\_RAT  
44 81.5 4.3 347 1 Y285\_MYCGE  
45 81.5 4.3 1385 1 CTAL\_MOUSE

## ALIGNMENTS







```

CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).
CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND SEC1) WHICH ENCODE
CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Y09882; CA71008.1; -.
CC EMBL: AF113532; AAD25351.1; -.
CC MGD: MGI:109374; Fut2.
CC InterPro: IPR002516; GT_11.
CC Pfam: PF01531; Glyco.transf.11; 1.
CC Signal-anchor; Golgi stack.
CC Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane;
CC Signal-anchor; Golgi stack.
CC DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 21 41 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC (POTENTIAL).
CC DOMAIN 42 368 LUMENAL, CATALYTIC (POTENTIAL).
CC CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 315 315 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 368 AA; 41464 MW; 4093B853EB37303B CRC64;
CC
CC Query Match 70.6%; Score 1338; DB 1; Length 368;
CC Best Local Similarity 88.0%; Pred. No. 9,9e-106;
CC Matches 243; Conservative 16; Mismatches 17; Indels 0; Gaps 0;
CC
Qy 46 GMFTINSTGRGNQMGVATLFAARMNGRIAFIPASMHNALAPIFRISLPVLHSDTAKK 105
Db 69 GMFTIRVGRGNQMGVATLFAARMNGRIAFIPASMHSTLAPIFRISLPVLHSDTAKR 128
Qy 106 IPWQNYHLNDWMEERYRHIPGHEVFTGPCSWTYHHLRPEILKEFTLHDHVRERAAQAF 165
Db 129 IPWQNYHLNDWMEERYRHIPGHEVFTGPCSWTYHHLRPEILKEFTLHDHVRERAAQAF 188
Qy 166 LRGLRVNGSQSPSTFGVHVVRGDDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPFVFW 225
Db 189 LRGLRVNGSQSPSTFGVHVVRGDDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPFVFW 248
Qy 226 TSNMGAWCREINASRGDVFAGNIGESPAKDFALLTQCNTHTMTGTGFGIWAAYLAGG 285
Db 249 TSDDMAWCKRSTASRGDVFAGNIGESPAKDFALLTQCNTHTMTGTGFGIWAAYLAGG 308
Qy 286 DTIYLANVTLPDSPPLKPKFAAFPEWVGIPADL 321
Db 309 DTIYLANVTQNSPHTVFKPEAAFLPEWVGIAADL 344
RESULT 5
FUT2_RABIT STANDARD; PRT; 354 AA.
ID FUT2_RABIT
AC Q10983;
DT 01-OCT-1996 (Rel. 34, Created)
DI 01-OCT-1996 (Rel. 34, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (Secretor blood
DE group alpha-2-fucosyltransferase) (GDP-L-fucose:beta-D-
DE galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)FT 2)
DE (Fucosyltransferase 2).
GN FUT2 OR SEC1 OR RFT-II.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

```

```

OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95238380; PubMed=7721792;
RA Hitoshi S., Kusunoki S., Kanazawa I., Tsuji S.;
RT "Molecular cloning and expression of two types of rabbit beta-
RL J. Biol. Chem. 270:8844-8850(1995).
CC -!- FUNCTION: CREATES A SOLUBLE PRECURSOR OLIGOSACCHARIDE FUC-ALPHA
CC (1,2)GALBETA- CALLED THE H ANTIGEN WHICH IS AN ESSENTIAL
CC SUBSTRATE FOR THE FINAL STEP IN THE SOLUBLE A AND B ANTIGEN
CC SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE THE SAME ACCEPTOR
CC SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -!- TISSUE SPECIFICITY: SALIVARY AND LACTATING MAMMARY GLANDS.
CC -!- MISCELLANEOUS: THERE ARE THREE GENES (FUT1, FUT2 AND FUT3) WHICH
CC ENCODE GALACTOSIDE 2-L-FUCOSYLTRANSFERASE IN RABBIT. THEY ARE
CC EXPRESSED IN A TISSUE-SPECIFIC MANNER.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X80225; CAA56512.1; -.
CC InterPro: IPR002516; GT_11.
CC Pfam: PF01533; Glyco.transf.11; 1.
CC Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane;
CC Signal-anchor; Golgi stack.
CC DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 23 43 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC (POTENTIAL).
CC DOMAIN 44 354 LUMENAL, CATALYTIC (POTENTIAL).
CC CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 317 317 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 354 AA; 40035 MW; 1E2B831F9DA6CCB4 CRC64;
CC
CC Query Match 68.1%; Score 1291; DB 1; Length 354;
CC Best Local Similarity 83.9%; Pred. No. 8,9e-102;
CC Matches 235; Conservative 20; Mismatches 25; Indels 0; Gaps 0;
CC
Qy 46 GMFTINSTGRGNQMGVATLFAARMNGRIAFIPASMHNALAPIFRISLPVLHSDTAKK 105
Db 71 GMFTINAMGRGNQMGVATLFAAKENGPRAYIPAOHSTLAPIFRISLPVLHSTASR 130
Qy 106 IPWQNYHLNDWMEERYRHIPGHEVFTGPCSWTYHHLRPEILKEFTLHDHVRERAAQAF 165
Db 131 VPWQNYHLNDWMEERYRHIPVYVRLTGYPCSWTYHHLRHEILREFTLHDHVRERAAQAF 190
Qy 166 LRGLRVNGSQSPSTFGVHVVRGDDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPFVFW 225
Db 191 LRGLRVNGSRPSTFGVHVVRGDDYVHVMPNWKGVVADRGYLEQALDWFRAPTAPYFVFW 250
Qy 226 TSNMGAWCREINASRGDVFAGNIGESPAKDFALLTQCNTHTMTGTGFGIWAAYLAGG 285
Db 251 TSNMGAWCREINASRGDVFAGNIGESPAKDFALLTQCNTHTMTGTGFGIWAAYLAGG 310
Qy 286 DTIYLANVTLPDSPPLKPKFAAFPEWVGIPADLSPLL 325
Db 311 DTIYLANVTAPDSPPLHVKFAAFPEWVGITANMGAL 350
RESULT 6

```









```

OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=94150718; PubMed=7906398;
RX Watson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Fayello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lighthouse J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders E., Showkeen R.,
RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Spratt J.,
RA Wheldrake P., Watson A., Weinstock L., Wilkinson-Spratt J.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RN [2]
RP REVISIONS.
RA Waterston R.; to the EMBL/GenBank/DBJ databases.
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: L16559; AAA27932.2; -
DR WormPep: C05E1.7; CE30483.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Hypothetical protein; Transferase; Glycosyltransferase.
SQ SEQUENCE 365 AA; 41991 MW; B5FBCA363F31977F CRC64;
-----
Query Match 8.9%; Score 168; DB 1; Length 365;
Best Local Similarity 22.0%; Pred. No. 7.9e-07;
Matches 71; Conservative 48; Mismatches 115; Indels 88; Gaps 16;
-----
QY 52 SIGRLGNMGGEYATLAFALARMNGR--LAFIPASMHNALAPTRISLPVLHSD----- 101
Dy 70 STARLANHIFELVSIVGMAKSLNKPALFVEDSKYNLLITGVKVLGLDLEFOIFEPV 129
QY 102 --TAKKIP-----WQNYHLNDWMEERYRHIPGHVFTGYPCSWFTYHHLRPEILKEF 152
Dy 130 HNKATKVPLEKCKCIFDNPDKFNISSELYHLTHGFYQ-----SWKYF----- 172
QY 153 TLHDHVEEAQAFRLGL-----RVNGSOPSTFVG---VHVRGDIV---HVMPNWKGVVA 202
Dy 173 ---DKYKEKVSFVKPAIDFSPSPNSDSNFISRICTHIRETDFDQGHSSNV----- 223
QY 203 DRGYLEKALDMFRAR-----YSSPFVFTVSGMAWCWRENINASRGD 243
Dy 224 --SPTKPALEFIKERQKQVKNKMLTVIMGDDPDPEAKMF-----EGTVRAKKEKIEET 277
QY 244 VVFAGNGIEGSPAKDFALL--TQCNIIMTI--GTFGIWAAYLAGSDTIYLANI--TLDPSP 299
Dy 278 KYFVS---ENTPQDLAYSHYSCDALLITAPSSFTGWLGLYLSKGQAVYYCDIRSTNDV 334
QY 300 FLK-VFKPEAFLPEWICPAD 320
Dy 335 YKKGVLDPDDDFVPSWTSIMLD 356
-----
RESULT 13

```

```

VGLD_HSVK
ID VGLD_HSVK STANDARD; PRT; 442 AA.
AC P22484;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein D precursor (glycoprotein 17/18).
GN GD OR GPI7/18 OR 72.
OS Equine herpesvirus type 1 (strain Kentucky A) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10329;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=91082407; PubMed=1845821;
RX Flowers C.C., Eastman E.M., O'Callaghan D.J.;
RA "Sequence analysis of a glycoprotein D gene homolog within the unique
RT short segment of the EHV-1 genome."
RL Virology 180:173-184(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=92263758; PubMed=1316673;
RX Colle C.F. III, Flowers C.C., O'Callaghan D.J.;
RA "Open reading frames encoding a protein kinase, homolog of
RT glycoprotein gx of pseudorabies virus, and a novel glycoprotein map
RT within the unique short segment of equine herpesvirus type 1."
RL Virology 188:545-557(1992).
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN D FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: M62923; AAA45081.1; -
DR EMBL: M86931; -; NOT_ANNOTATED_CDS.
DR EMBL: M87497; AAA46073.1; ALT_INIT.
DR PIR: A38518; VGBEEA.
DR InterPro: IPR002896; Herpes_glycop_D.
DR Pfam: PF01537; Herpes_glycop_D; 1.
KW Glycoprotein; Signal; Transmembrane.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 442 GLYCOPROTEIN D.
FT DOMAIN 20 405 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 406 422 POTENTIAL.
FT DOMAIN 423 442 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 103 103 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 111 111 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 347 347 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 396 396 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 442 AA; 49908 MW; 323CDCA9C9762F05 CRC64;
-----
Query Match 4.8%; Score 91; DB 1; Length 442;
Best Local Similarity 22.7%; Pred. No. 2.9;
Matches 54; Conservative 31; Mismatches 103; Indels 50; Gaps 10;
-----
QY 1 LQORIVKIQ-----PLSEKELPMTTOMSSCNTSESPENRRSEGHGNGELGMEI 50
Dy 234 LYRRVIEDGRRIVTDFSVTPSERCPTAFELNFCN---PORCKTPEQYSRGE---VTR 287
QY 51 NSIGRLGNMGGEYAT--LAFALARMNGRLAFIPASMHNALAPTRISLPVLHSDTAKKIPW 108
Dy 288 RFLGLENPFQGEHMTWVKFWFYDGNL--PVQFYEAQA-----FAPYVZP 331
QY 109 QNYHLNDWMEERYRH-----IPGHVFTGYPCSWFTYHHLRPEILKEFTLHDHVRDAQ 163
Dy 332 DNHGCFDSVESEITQNKTDPKQADPKPNQPFKWPISKHLVPRD-----DEVDEVIE 384
QY 164 AFLRLGVNGSQPSTFVGHVHVRGDYVHVMNPNVKGVVADRGYLEKALDMFRARYSSP 222

```

Db 385 PVTPEKTSKN-STFVGISVGLGIAGLVGVILYVCLRR---KKELKVCTERLDSP 438

RESULT 14

FMA3\_FORGI STANDARD; PRT; 353 AA.

AC Q51826;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Major fimbrial subunit protein, type III precursor (Fimbrillin)

DE (Fimbrillin).

GN FIMA.

OS Porphyromonas gingivalis (Bacteroides gingivalis).

OC Bacteria; Bacteroidetes; Bacteroides; Bacteroidales;

OC Porphyromonadaceae; Porphyromonas.

OX NCBI\_TaxID=837;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=6/26;

RX MEDLINE=94071950; PubMed=7902712;

RA Fujiwara T., Morishima S., Takahashi I., Hamada S.;

RT "Molecular Cloning and sequencing of the fimbrillin gene of

RT Porphyromonas gingivalis strains and characterization of recombinant

RT proteins.";

RL Biochem. Biophys. Res. Commun. 197:241-247(1993).

RN [2]

RP FUNCTION, AND CLASSIFICATION INTO TYPES.

RX MEDLINE=21614934; PubMed=11748193;

RA Nakagawa I., Amano A., Kuboniwa N., Nakamura T., Kawabata S.,

RA Hamada S.;

RT "Functional differences among FimA variants of Porphyromonas

RT gingivalis and their effects on adhesion to and invasion of human

RT epithelial cells.";

RL Infect. Immun. 70:277-285(2002).

CC -!- FUNCTION: Fimbrillin is the structural subunit of the fimbriae,

CC that are filamentous appendages on the cell surface. Fimbriae of

CC P.gingivalis are recognized as a major virulence factor as they

CC mediate cell adhesion and play an important role in invasion of

CC periodontal tissues.

CC -!- SUBCELLULAR LOCATION: Fimbria.

CC -!- SIMILARITY: BELONGS TO THE P.GINGIVALIS FIMBRILLIN FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; D17801; BAA04627.1; -

KW Fimbria; Virulence.

FT PROPEP 1 10 BY SIMILARITY.

FT CHAIN 11 353 MAJOR FIMBRIAL SUBUNIT PROTEIN, TYPE III.

SO SEQUENCE 353 AA; 38024 MW; 7FBE4FBF427EA2AB CRC64;

Query Match 4.7%; Score 89; DB 1; Length 353;

Best Local Similarity 22.1%; Pred. No. 3.3;

Matches 64; Conservative 38; Mismatches 110; Indels 78; Gaps 12;

Qy 18 MTTOMSSNTSPEMRRDSEQ-----HGNGELRGMTFINSIGRLGNMG---EYATLFA 68

Db 80 LTTTELGNEQAAGLIMTAEPVEVFLVAGN-----NYGYDGSGGNGISGCTPLE 130

Qy 69 LARMNGRLAF--IPASMHNALAPIRISLPVLSHTAKK----- 105

Db 131 IKRVHARIAFTKIEVMSQSVANKNFAPENIYALVAKKSNLFGASLNSDDAYLTGSL 190

Qy 106 -----IPQNYHLNDWEERYRHI-----PGHFVFTGYPCSWTFYHH--LRPEIL- 149

Db 191 ITFNAGYSPANYTHVDNLGROYTEGATVNTPKGF-----YVLESTYQAQNLARPTILC 245

Qy 150 --KBEFTLHDHV---REEAQAFRLGLRVNGSQPSTF--VGHVRRGDYVHVPNWKGVV 201

Db 246 VKGKLTAKHGDGALSSEMTAAFNAGWIVANNDPTYYPYLVNFESNNYTYTGEAVEKSKI 305

Qy 202 ADRGYLEKALDMFRARYSSPVFVTSNGMARENINASRGDVVFAGNGI 251

Db 306 VRNHKFDINLTITGFTNNPENFIT-----ESANLNVNCVAAWKGV 347

RESULT 15

CDH\_PHACH STANDARD; PRT; 773 AA.

AC Q01738; C00047;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cellulose dehydrogenase precursor (EC 1.1.5.1) (CDH) (Cellulobiose-

DE guinone oxidoreductase).

GN CDH-1 AND CDH-2.

OS Phanerochaete chrysosporium.

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;

OC Aphyllophorales; Corticiaceae; Phanerochaete.

OX NCBI\_TaxID=5306;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OGC101;

RX MEDLINE=97077226; PubMed=8919793;

RA Li B., Nagalla S.R., Renganathan V.;

RT "Cloning of a cDNA encoding cellulobiose dehydrogenase, a

RT hemoflavoenzyme from Phanerochaete chrysosporium.";

RL Appl. Environ. Microbiol. 62:1329-1335(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=OGC101;

RX MEDLINE=97176414; PubMed=9023960;

RA Li B., Nagalla S.R., Renganathan V.;

RT "Cellulobiose dehydrogenase from Phanerochaete chrysosporium is encoded

RT by two allelic variants.";

RL Appl. Environ. Microbiol. 63:796-799(1997).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 19-208.

RX MEDLINE=20139694; PubMed=10673428;

RA Hallberg B.M., Bergfors T., Boeckbro K., Pettersson G., Henriksson G.,

RA Divine C.;

RT "A new scaffold for binding haem in the cytochrome domain of the

RT extracellular flavocytochrome cellobiose dehydrogenase.";

RL Structure 8:79-88(2000).

CC -!- FUNCTION: DEGRADATES BOTH LIGNIN AND CELLULOSE. OXIDIZES CELLOBIOSE

CC TO CELLOBIONOLACTONE.

CC -!- CATALYTIC ACTIVITY: Cellobiose + a quinone = cellobiono-1,5-

CC lactone + a phenol.

CC -!- COFACTOR: ONE FAD AND ONE HEME B.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE GMC

CC OXIDOREDUCTASES FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; U46081; AAC49277.1; -

DR EMBL; U65888; AB61455.1; -

DR EMBL; U50409; AB92262.1; -

DR PDB; 1D7B; 18-OCT-99.

DR PDB; 1D7C; 18-OCT-99.

DR PDB; 1D7D; 18-OCT-99.

DR InterPro; IPR000172; GMC\_oxred.

DR InterPro; IPR001100; Pyr\_redox.

DR Pfam; PF00732; GMC\_oxred; 1.

```
DR PRINTS: PRO0411; PNDRPTASEI.
DR PROSITE: PS00623; GMC_OXRED_1; 1.
DR PROSITE: PS00624; GMC_OXRED_2; 1.
KW Cellulose degradation; Oxidoreductase; FAD; Flavoprotein; Heme;
KW Multigene family; Signal; 3D-structure.
FT SIGNAL 1 18
FT CHAIN 19 773 CELLOBIOSE DEHYDROGENASE.
FT DOMAIN 19 238 HEME DOMAIN.
FT DOMAIN 235 773 OXIDOREDUCTASE.
FT BINDING 83 83 HEME LIGAND.
FT BINDING 181 181 HEME LIGAND.
FT NP_BIND 236 265 FAD (ADP PART) (POTENTIAL).
SQ SEQUENCE 773 AA: 32007 MW; 54F721E779AA4D7B CRC64;

Query Match      4.7%; Score 89; DB 1; Length 773;
Best Local Similarity 22.0%; Pred. No. 8.5;
Matches 49; Conservative 29; Mismatches 91; Indels 54; Gaps 11;

QY 126 GHVRETGYPCSWTFYHHLRPETLKEFTLHDHYREAAQAFLRG-----LEVNGSQPS 177
Db 337 GDESSVGNPSSWNTNHPYTSKLSRLPSTDHPSTDGQRYLEQSFNVVSQLKQGQYNQA 396
QY 178 TFVGVHVRGQDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVWTSNGMAWCRENI 237
Db 397 T-----INDNPYKOHVF---GY--SAFDFLNGKRGAPVATYLTALA--RPNF 438
QY 238 ----NASRGDVVFAGNGIEGSPAKD-----FALLTQCQNHITMTIGTFG----IWAAYL 282
Db 439 TEKTNVNVSVNVRNGSQLGVQNDPTLGNFGFIPVTPKGRVILSAGAFGTSKILFQSGI 498
QY 283 AGGDTIYLANYLDPDSPFLKVPKPEAAFLP---EWVGIPADLS 322
Db 499 GPTDMI-----QTVQSNP-----TAAALPPQNWNLNLPVGMN 531

Search completed: May 27, 2003, 15:07:14
Job time : 11.3317 secs
```



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:52:27 ; Search time 43.6293 Seconds  
(without alignments)  
1050.630 Million cell updates/sec

Title: US-10-040-863-11

Perfect score: 1850

Sequence: 1 MLVQMPSPFMAHILFV.....AAFLPEWTGIAADLSPLLKH 344

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002:\*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1850	100.0	344	21	Human Sec2 catalyt
2	1832.5	99.1	343	19	Human Sec2 protein
3	1528	82.6	340	19	Porcine secretor t
4	1528	82.6	340	19	Pig secretor. Sus
5	1475	79.7	380	21	Rat hepatoma H35 c
6	1475	79.7	380	23	Rat hepatoma H35 c
7	1376	74.4	353	21	Rat hepatoma H35 c
8	1376	74.4	353	23	Rat hepatoma H35 c
9	1081.5	58.5	365	20	Swine alpha(1,2) f
10	1081.5	58.5	365	20	Swine alpha-1,2-fu

11	1080.5	58.4	365	12	AA13751
12	1080.5	58.4	365	16	AA180154
13	1080.5	58.4	365	16	AA170422
14	1080.5	58.4	365	16	AA170421
15	1080.5	58.4	365	17	AA190572
16	1080.5	58.4	365	18	AA123805
17	1080.5	58.4	365	18	AA113640
18	1080.5	58.4	365	21	AA197279
19	1075.5	58.1	365	19	AA153102
20	1075.5	58.1	365	21	AA179302
21	1075.5	58.1	365	23	AA147995
22	1072.5	58.0	365	15	AA145936
23	987.5	53.4	357	20	AA117969
24	532	28.8	110	21	AA103530
25	148	8.0	287	23	AA151992
26	144.5	7.8	300	21	AA192713
27	144.5	7.8	300	22	AA164071
28	100	5.4	355	21	AA170408
29	92	5.0	575	18	AA122124
30	91	4.9	575	23	AA134136
31	90.5	4.9	661	22	AA139722
32	90	4.9	515	22	AA139884
33	90	4.9	533	12	AA115057
34	90	4.9	575	18	AA122125
35	90	4.9	575	23	AA108405
36	90	4.9	1073	22	AA128599
37	88.5	4.8	417	22	AA179361
38	87.5	4.7	365	19	AA180571
39	87.5	4.7	397	19	AA180573
40	87	4.7	637	22	AA171200
41	86.5	4.7	417	22	AA178377
42	86.5	4.7	958	21	AA151120
43	85.5	4.6	547	22	AA124778
44	85.5	4.6	739	22	AA193019
45	85	4.6	402	22	AA1863790

#### ALIGNMENTS

RESULT 1  
AAB36106  
ID AAB36106 standard; Protein; 344 AA.

AC AAB36106;

DT 19-FEB-2001 (first entry)

DE Human Sec2 catalytic domain.

Human; Sec2; alphas-2fucosyltransferase; cytostatic;

neuroprotective; nootropic; gene therapy; Fucalpal-2Galbeta1-3GalNAc;

immunotherapy; immunosuppression; cancer; neurological disease;

small cell lung carcinoma.

OS Homo sapiens.

PN WO200064464-A1.

PD 02-NOV-2000.

PF 23-APR-1999; 99WO-US07384.

PR 23-APR-1999; 99WO-US07384.

(PACI-) PACIFIC NORTHWEST CANCER FOUND.

Holmes EH, Sherwood AL;

WPI; 2000-687262/67.

New rat ganglioside GM1-specific alpha-2fucosyltransferase, useful for

preparation of fucosyl GM1 which is useful as a nutritional composition

PT or immunotherapeutic for cancer and neurological diseases -

PS Example; Fig 3B; 91pp; English.

XX The present sequence is given in a specification relating to a rat  
CC ganglioside GM1-specific alpha(1,2)-fucosyltransferase protein. The protein  
CC or its cellular fraction is useful for synthesis of a molecule comprising  
CC Fucal-2galactose-3galNAc, a glycolipid, glycoprotein, glycolipoprotein  
CC or a free oligosaccharide comprising Fucal-2galactose-3galNAc.  
CC The method involves contacting alpha(1,2)-fucosyltransferase with GDP-fucose  
CC and a molecule or glycolipid, glycoprotein, glycolipoprotein or  
CC oligosaccharide having a terminal galactose-3galNAc group. It is also  
CC useful for synthesis of fucosyl-GM1 by contacting the protein with  
CC GDP-fucose and ganglioside GM1. The obtained glycoproteins,  
CC glycolipoproteins, glycolipids and oligosaccharides are useful as  
CC nutritional compositions and fucosyl-GM1 is useful for inducing an  
CC immunotherapeutic or immunosuppressive action against cancer,  
CC neurological disease or small cell lung carcinoma.

XX Sequence 344 AA;

Query Match 100.0%; Score 1850; DB 21; Length 344;  
Best Local Similarity 100.0%; Pred. No. 2.5e-186;  
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLVYMPFSPMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSQL 60  
DB 1 MLVYMPFSPMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSQL 60  
QY 61 RGMWTINAIGRLGNQMGEXATLYALAKMNGRPAPFPAQMHSSTLAPFIPRITLPVLHSATAS 120  
DB 61 RGMWTINAIGRLGNQMGEXATLYALAKMNGRPAPFPAQMHSSTLAPFIPRITLPVLHSATAS 120  
QY 121 RHPWNYHLNDWMEEEYRHIPPGEYVFTGYPSCWTFYHHLRQELQFTLHDHVRREAAQ 180  
DB 121 RHPWNYHLNDWMEEEYRHIPPGEYVFTGYPSCWTFYHHLRQELQFTLHDHVRREAAQ 180  
QY 181 KFLRGLQVNSRPGTFVGVHVVRGDIYVHVPKVKGVVADRRYLOQALDWFPRARYSSLIIF 240  
DB 181 KFLRGLQVNSRPGTFVGVHVVRGDIYVHVPKVKGVVADRRYLOQALDWFPRARYSSLIIF 240  
QY 241 VVTSNGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTGIWAAAYLT 300  
DB 241 VVTSNGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTGIWAAAYLT 300  
QY 301 GGDITILANYTLDPSPFLKIFKPEAAFLPEWTGIAADLSPLLKH 344  
DB 301 GGDITILANYTLDPSPFLKIFKPEAAFLPEWTGIAADLSPLLKH 344

RESULT 2  
AAW69332  
ID AAW69332 standard; Protein; 343 AA.

XX AAW69332;

XX 20-NOV-1998 (first entry)

DE Human Sec2 protein sequence.

XX Sec2; alpha(1,2) fucosyltransferase; H blood group; secretor genotyping;  
KW GDP-L-fucose; beta-D-galactoside 2-alpha-L-fucosyltransferase; human;  
KW FUT2; nonsecretor genotyping.

XX Homo sapiens.

XX US5807732-A.

XX 15-SEP-1998.

XX 28-FEB-1995; 95US-0395800.

XX 28-FEB-1995; 95US-0395800.

PR

XX (GIOR/) GIORGI D.  
PA (KELL/) KELLY R.J.  
PA (LENN/) LENNON G.  
PA (LOWE/) LOWE J.B.  
PA (ROUQ/) ROQUIER S.

XX Giorgi D, Kelly RJ, Lennon G, Lowe JB, Rouquier S;  
PI WPI: 1998-520127/44.  
XX DR N-PSDB; AAV58323.

XX DNA encoding fucosyltransferase enzyme - useful for producing  
PT recombinant enzyme and genotyping person as secretor or nonsecretor  
PT Disclosure; Column 45-50; 55pp; English.

XX This sequence is the human Sec2 protein of the invention. The DNA  
CC encodes a alpha(1,2) fucosyltransferase and is the Secretor  
CC alpha(1,2) fucosyltransferase locus, that cross hybridises with the  
CC H blood group alpha(1,2) fucosyltransferase gene. The DNA is useful for  
CC producing a recombinant human GDP-L-fucose:beta-D-galactoside  
CC 2-alpha-L-fucosyltransferase (FUT2) which can be used for genotyping an  
CC individual as a secretor or nonsecretor as it is known that nonsecretors  
CC homozygous for a mutant allele of the FUT2 gene that has a stop codon in  
CC the position corresponding to amino acid 143.

XX Sequence 343 AA;

Query Match 99.1%; Score 1832.5; DB 19; Length 343;  
Best Local Similarity 99.7%; Pred. No. 1.8e-184;  
Matches 343; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MLVYMPFSPMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSQL 60  
DB 1 MLVYMPFSPMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSQL 60  
QY 61 RGMWTINAIGRLGNQMGEXATLYALAKMNGRPAPFPAQMHSSTLAPFIPRITLPVLHSATAS 120  
DB 61 RGMWTINAIGRLGNQMGEXATLYALAKMNGRPAPFPAQMHSSTLAPFIPRITLPVLHSATAS 120  
QY 121 RHPWNYHLNDWMEEEYRHIPPGEYVFTGYPSCWTFYHHLRQELQFTLHDHVRREAAQ 180  
DB 121 RHPWNYHLNDWMEEEYRHIPPGEYVFTGYPSCWTFYHHLRQELQFTLHDHVRREAAQ 180  
QY 181 KFLRGLQVNSRPGTFVGVHVVRGDIYVHVPKVKGVVADRRYLOQALDWFPRARYSSLIIF 240  
DB 181 KFLRGLQVNSRPGTFVGVHVVRGDIYVHVPKVKGVVADRRYLOQALDWFPRARYSSLIIF 240  
QY 241 VVTSNGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTGIWAAAYLT 300  
DB 241 VVTSNGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTGIWAAAYLT 300  
QY 301 GGDITILANYTLDPSPFLKIFKPEAAFLPEWTGIAADLSPLLKH 344  
DB 301 GGDITILANYTLDPSPFLKIFKPEAAFLPEWTGIAADLSPLLKH 344

RESULT 3

AAW37855  
ID AAW37855 standard; Protein; 340 AA.

XX AAW37855;

XX 28-AUG-1998 (first entry)

XX Porcine secretor transferase (FUT2).

XX Secretor; glycosyltransferase; FUT2; pig; epitope; antigen;  
KW transgenic animal; xenotransplantation; organ transplant.

XX Sus scrofa.

XX

```

FH Key Location/Qualifiers
FT Domain 1..4
FT /note= "N-terminal cytoplasmic tail"
FT Domain 5..26
FT /note= "transmembrane domain"
FT Domain 27..340
FT /note= "C-terminal domain"
FT Modified-site 185..187
FT /note= "Asn is N-glycosylated"
FT Modified-site 251..253
FT /note= "Asn is N-glycosylated"
FT Modified-site 279..281
FT /note= "Asn is N-glycosylated"
FT Modified-site 305..307
FT /note= "Asn is N-glycosylated"
XX
XX WO9807837-A1.
XX
XX 26-FEB-1998.
XX
XX 22-AUG-1997; 97WO-AU00540.
XX
XX 23-AUG-1996; 96AU-0001823.
XX
XX (AUST-) AUSTIN RES INST.
XX
XX McKenzie IFC, Sandrin MS;
XX
XX WPI: 1998-169148/15.
XX N-PSDB; AAV29003.
XX
XX Nucleic acid encoding glycosyltransferase able to compete with
XX second such enzyme - particularly used to reduce expression of
XX unwanted carbohydrate epitope(s) on tissues intended for
XX transplantation
XX
XX Claim 6; Fig 1A-B; 40pp; English.
XX
XX This polypeptide comprises porcine secretor glycosyltransferase
XX (SE or FUT2), a type II integral membrane protein has high affinity
XX for type I and type III substrates. Its amino acid sequence was
XX deduced from the nucleotide sequence of a genomic DNA clone (see
XX AAV29003) isolated from a pig liver library on the basis of homology
XX to human FUT2 cDNA. Pig FUT2 shows 83.2% identity with human FUT2.
XX Expression of a glycosyltransferase, such as porcine Se, in a
XX tissue results in reduced expression of unwanted carbohydrate
XX epitopes on the tissue, especially porcine heart, liver, kidney or
XX pancreas, rendering it more suitable for transplantation, i.e. less
XX immunogenic and of increased immunological acceptability. A
XX claimed method of producing a cell from a donor species that is
XX immunologically acceptable to a recipient species involves reducing
XX levels of carbohydrate on the donor cell that causes it to be
XX recognised as non-self by the recipient by expressing a nucleic
XX acid for a glycosyltransferase such as porcine Se in the cell.
XX
XX Sequence 340 AA;
XX
XX Query Match 82.6%; Score 1528; DB 19; Length 340;
XX Best Local Similarity 82.0%; Pred. No. 2.3e-152;
XX Matches 282; Conservative 24; Mismatches 34; Indels 4; Gaps 2;
XX
XX 1 MLVQMPSFPMHAFILFVFTYSTIFHVQORLAKTCAMWELPVQIPVLASTSKALGPSOL 60
XX || : || || || || || || || || || || || || || || || || || ||
XX 1 MLSMQASFFPTGPPILFVFTASTIFHLQQRWVKIQPTWELQM---VTQVTESPSPOL 57
XX
XX 61 RGMWITNATIGRLGNMGVEYATLYALAKNGRFAFIPAOHMSHTLAFITFLPVLHSATAS 120
XX : || || || || || || || || || || || || || || || || || || ||
XX 58 KGMWITNATIGRLGNMGVEYATLYALAKNGRFAFIPPEHMSHTLAFITFLPVLHASTAR 117
XX
XX 121 RIPWQNYHLNDWMEERYRHI-PGEVVRITGYPCSWTFYHHLRQELIQPFTLHDHVREPAQ 180
XX || || || || || || || || || || || || || || || || || || ||
XX 118 RIPWQNYHLNDWMEERYRHI-PGEVVRITGYPCSWTFYHHLRQELIQPFTLHDHVREPAQ 176

```

```

QY 181 KFLRGLQVNGSRPGTFVGVHVRGQDVHVMPKVKWGVADRRYVQQALDWFRRARYSSLIF 240
DB 177 DELRGLRVNGSRPSTYGVGVHVRGQDVHVMPKVKWGVADRRYVQQALDWFRRARYSPVF 236
QY 241 VVTSNGMWCNRNIDTSHGDVVFAGDGISSPAKDFALLTQCNHTTWTCTGFWAAYLT 300
DB 237 VVSSNGMWCNRNINASGDVVFAGDGISSPAKDFALLTQCNHTTWTCTGFWAAYLA 296
QY 301 GGDTIYLIANTLPDPSFLKIFKPEAAFLPEWTGIAADLSPLLK 344
DB 297 GGDTIYLIANTLPDPSFLKIFKPEAAFLPEWTGIEADLSPLLK 340
XX
XX AAW53101
XX ID AAW53101 standard; Protein; 340 AA.
XX
XX AC AAW53101;
XX
XX DT 08-JUL-1998 (first entry)
XX
XX DE Pig secretor.
XX
XX KW Pig; secretor; chimeric; glycosyltransferase; gene therapy;
XX KW transplantation.
XX
XX OS Sus scrofa.
XX
XX PN WO9805768-A1.
XX
XX PD 12-FEB-1998.
XX
XX PF 01-AUG-1997; 97WO-AU00492.
XX
XX PR 21-AUG-1996; 96US-0024279.
XX PR 02-AUG-1996; 96AU-0001402.
XX
XX PA (AUST-) AUSTIN RES INST.
XX
XX PI McKenzie IFC, Sandrin MS;
XX
XX XX WPI: 1998-159170/14.
XX DR N-PSDB; AAV21639.
XX
XX Nucleic acids encoding chimeric glycosyltransferases - used for
XX altering carbohydrate levels on the surface of cells, useful in gene
XX therapy and transplantation
XX
XX Example 2; Fig 6; 51pp; English.
XX
XX The present sequence represents pig secretor used in an example of the
XX present invention. The present invention describes nucleic acids (NA)
XX encoding a chimeric glycosyltransferase. The NAs comprise a catalytic
XX domain of a first glycosyltransferase (GT) and a localisation signal of
XX a second GT, whereby when the NA is expressed in a cell and where the
XX chimeric enzyme is located in an area of the cell where it is able to
XX compete for substrate with a second GT, resulting in reduced levels of
XX a product from the second GT. The NAs can be used to produce cells and
XX organs with desired glycosylation patterns. Products and methods of the
XX present invention can be used to reduce the levels of undesirable
XX epitopes in cells, tissues or organs which may be used in
XX CC transplantation or gene therapy.
XX
XX Sequence 340 AA;
XX
XX Query Match 82.6%; Score 1528; DB 19; Length 340;
XX Best Local Similarity 82.0%; Pred. No. 2.3e-152;
XX Matches 282; Conservative 24; Mismatches 34; Indels 4; Gaps 2;
XX
XX 1 MLVQMPSFPMHAFILFVFTYSTIFHVQORLAKTCAMWELPVQIPVLASTSKALGPSOL 60
XX || : || || || || || || || || || || || || || || || || || ||
XX 1 MLSMQASFFPTGPPILFVFTASTIFHLQQRWVKIQPTWELQM---VTQVTESPSPOL 57

```

XX	Sequence	380 AA;
Qy	Query Match	79.7%; Score 1475; DB 21; Length 380;
Db	Best Local Similarity	77.7%; Pred. No. 1.le-146;
Matches	275; Conservative	30; Mismatches 37; Indels 12; Gaps
Qy	1 MLVQMPFSPFMAHFLIFVFTVSIFHVQQRKAKIQAM--WELPV-----QIEVLA 49	
Db	1 MASAQVFPSPFLAHFLIFVFTSIIHLQORIVKLQPLSEKELPNTTQMSSGNTESPEMR 60	
Qy	50 STSKALGPSQRGQWMTINAGRLGNQMGVEATLYALAKMNGRPAFIPQMHSTLAPIFRI 109	
Db	61 RDSQHGNGELRGFTTINSIGRLGNQMGVEATLFAALMNGRLAFIPASMHNALAPIFRI 120	
Qy	110 TLPVLHSATASRIQWNYHLNDWMEERYRHIPGCEYVRFTGYPCSWTFYHHLROEILQEF 169	
Db	121 SLPLVHSDTAKKIQWNYHLNDWMEERYRHH--PGHFVRFTGYPCSWTFYHHLROEILKEF 179	
Qy	170 TLHDVREEAQKFLRGQLVNGSRFGTIVGVHVRGRGDVHVMPKVKGWADRRRYLQALD 229	
Db	180 TLHDVREEAQKFLRGVLVNGSQPSTFVGHVVRGRGDVHVMPNVKGVADRGYLEKALD 239	
Qy	230 WFRARYSLFVVTSGMWCARENIDTSHGSDVVFAGDLEGSPAKDFALLTQCNTHTMTI 289	
Db	240 MFRARYSPFVVTSGMWCARENINASRGDVPVAGNGIEGSPAKDFALLTQCNTHTMTI 299	
Qy	290 GTFGIWAAYLTGGDTIYLIANTLTDPSPFLKIFKPEAAFLPEWTGCIADLSPLLK 343	
Db	300 GTFGIWAAYLAGGDTIYLIANTLTDPSPFLKVFKEAAFLPEWVGIPADLSPLLK 353	
RESULT 6		
AAE16622	AAE16622 standard; Protein; 380 AA.	
XX	AC	AAE16622;
XX	DT	09-APR-2002 (first entry)
XX	DE	Rat hepatoma H35 cell alphas-2FucT.
XX	KW	Rat: alphas-2Fucosyltransferase; alphas-2FucT; antisense therapy;
XX	KW	galactose beta-1,3N-acetylgalactosamine; Galbeta1-3GalNAc; glycolipid;
XX	KW	glycoprotein; glycolipoprotein; oligosaccharide; fucosyl-GM1; cancer;
XX	KW	gene therapy; oncogenic transformation; cytostatic; ganglioside; GM1;
XX	KW	cell transformation.
XX	OS	Rattus norvegicus.
XX	PH	Key Location/Qualifiers
FT	Domain	1..27 "Intracellular/Transmembrane domain"
FT	Modified-site	/note= "N-glycosylated"
FT	Region	221..380
FT	Modified-site	/note= "N-glycosylated"
FT	Modified-site	/note= "N-glycosylated"
FT	Modified-site	/note= "N-glycosylated"
PN	US6329170-B1.	
XX	11-DEC-2001.	
XX	23-APR-1999;	99US-0298886.
XX	23-APR-1999;	99US-0298886.
XX	(WHO-) NORTHWEST HOSPITAL.	
XX	Holmes EH.	Sherwood AL;

```

XX WPI: 2002-121132/16.
DR N-PSDB: AAC27207.
XX
XX Rat hepatoma H35 cell alpha1-2fucosyltransferase, useful for producing
PI GM1-specific alpha1-2fucosyltransferase enzyme by recombinant
PT techniques and for detecting oncogenic transformation of test tissues -
XX
XX Claim 3; Fig 5; 41pp; English.
PS
XX The invention relates to rat GM1-specific alpha1-2fucosyltransferase
XX (alpha1-2FucT) enzyme and its corresponding nucleic acid. This nucleic
CC acid is specific for a carbohydrate moiety found in ganglioside GM1,
CC a terminal galactose beta1-3N-acetylgalactosamine (Galbeta1-3GalNAc)
CC saccharide. Alpha1-2FucT DNA is useful for producing rat alpha1-2FucT
CC protein by recombinant techniques. Alpha1-2FucT DNA is useful for the
CC preparative synthesis of fucosyl containing glycolipids, glycoproteins,
CC glycolipoproteins and oligosaccharide, and for preparing fucosyl-GM1.
CC Alpha1-2FucT DNA is useful for detecting oncogenic transformation which
CC involves assaying for changes in expression of alpha1-2 FucT. Since
CC alpha1-2FucT is activated in cell transformation, antisense sequences
CC derived from alpha1-2FucT DNA are useful for inhibiting, suppressing
CC or treating cancer. Alpha1-2FucT DNA is useful in gene therapy and
CC antisense therapy. The present sequence is rat hepatoma H35 cell
CC alpha1-2FucT.
XX
XX Sequence 380 AA;
SQ
Query Match 79.7%; Score 1475; DB 23; Length 380;
Best Local Similarity 77.7%; Pred. No. 1.1e-146;
Matches 275; Conservative 30; Mismatches 37; Indels 12; Gaps 3;
QY 1 MLVQPFSPMAHFLFVTVTITHVQRLAKIQAM--WELPV-----QIPVLA 49
Db 1 MASQVFPSPPLAHFLFVFTVTHLQRIYKLOPLSEKELPMTQMSSGNTESPEMR 60
QY 50 STSKALGPSLGRMTNATIGRLGNOMGEVATLYALAKMNGRPAFIPAQMHSFLAPFRI 109
Db 61 RDEQHGNGELRGWFTNSIGRLGNOMGEVATLFAARMNGRLAFIPASMHNALAPFRI 120
QY 110 TLPVLHSATASRIPWQNYHLNDWMEERYHIPPGEYVRFYGYPCSWTFYHHLRQEILOEF 169
Db 121 SLEVLHSDTAKKIPWQNYHLNDWMEERYHRI-PGHFVRFYGYPCSWTFYHHLRPEILKEF 179
QY 170 TLHDHVREREAQKELRGLOVNGSRPGTFVGVHVRGDYVHVMPKVGWADRRYLOQALD 229
Db 180 TLHDHVREREAQKELRGLOVNGSQPSTFVGVHVRGDYVHVMPNWKGVADRGYLEKALD 239
QY 230 WFRARYSSLIFFVTSNGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTHTMTI 289
Db 240 MERARYSSPVFVTSNGMAWCRENINASRGDVVFAGNIEGSPAKDFALLTQCNTHTMTI 299
QY 290 GTFGIWAAYLTGGDTIYLANYTLPDSPFLKIFKPEAAFLPEWTGIAADLSPLLK 343
Db 300 GTFGIWAAYLAGDGTIYLANYTLPDSPFLKVPKPEAAFLPEWVGIPADLSPLLK 353
RESULT 7
AAB36105
ID AAB36105 standard; Protein; 353 AA.
XX
XX AAB36105;
XX
XX 19-FEB-2001 (first entry)
XX
XX Rat hepatoma H35 cell alpha1-2fucosyltransferase catalytic domain.
XX
XX Rat: alpha1-2fucosyltransferase; cytostatic; neuroprotective;
XX nootropic; gene therapy; Fucalpal-2Galbeta1-3GalNAc; immunotherapy;
XX immunosuppression; cancer; neurological disease.
XX small cell lung carcinoma.
XX
XX Rattus norvegicus.
OS

```

```

XX WO200064464-A1.
XX
XX 02-NOV-2000.
XX
XX 23-APR-1999; 99WO-US07384.
XX
XX 23-APR-1999; 99WO-US07384.
XX
XX (PACI-) PACIFIC NORTHWEST CANCER FOUND.
XX
XX Holmes EH, Sherwood AL;
XX
XX WPI: 2000-687262/67.
XX N-PSDB: AAC67966.
XX
XX New rat ganglioside GM1-specific alpha1-2fucosyltransferase, useful for
PT preparation of fucosyl GM1 which is useful as a nutritional composition
PT or immunotherapeutic for cancer and neurological diseases -
XX
XX Claim 2; Fig 3A; 91pp; English.
PS
XX The present sequence is given in a specification relating to a rat
XX ganglioside GM1-specific alpha1-2fucosyltransferase protein. The protein
CC or its cellular fraction is useful for synthesis of a molecule comprising
CC Fucalpal-2Galbeta1-3GalNAc, a glycolipid, glycoprotein, glycolipoprotein
CC or a free oligosaccharide comprising Fucalpal-2Galbeta1-3GalNAc.
CC The method involves contacting alpha1-2fucosyltransferase with GDP-fucose
CC and a molecule or glycolipid, glycoprotein, glycolipoprotein or
CC oligosaccharide having a terminal galbeta1-3GalNAc group. It is also
CC useful for synthesis of fucosyl-GM1 by contacting the protein with
CC GDP-fucose and ganglioside GM1. The obtained glycoproteins,
CC glycolipoproteins, glycolipids and oligosaccharides are useful as
CC nutritional compositions and fucosyl-GM1 is useful for inducing an
CC immunotherapeutic or immunosuppressive action against cancer,
CC neurological disease or small cell lung carcinoma.
XX
XX Sequence 353 AA;
SQ
Query Match 74.4%; Score 1376; DB 21; Length 353;
Best Local Similarity 78.9%; Pred. No. 2.7e-136;
Matches 288; Conservative 26; Mismatches 31; Indels 12; Gaps 3;
QY 28 VQORLAKIQAM--WELPV-----QIPVLASTSKALGPSLGRMTNATIGRLGNOM 76
Db 1 LQORIVKLOPLSEKELPMTQMSSGNTESPEMRDSDHNGELRGMTFINSIGRLGNOM 60
QY 77 GEVATLYALAKMNGRPAFIPAQMHSFLAPFRIITLPLVLSATASRIPWQNYHLNDWMEER 136
Db 61 GEVATLFAARMNGRLAFIPASMHNALAPFRIISLPVLHSDTAKKIPWQNYHLNDWMEER 120
QY 137 YRHIPPGEYVRFYGYPCSWTFYHHLRQEILOEFTHDVREREAQKELRGLOVNGSRPGTF 196
Db 121 YRHI-PGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREREAQKELRGLOVNGSQPSTF 179
QY 197 VGVHVRGDYVHVMPKVGWADRRYLOQALDWMFRARYSSLIFFVTSNGMAWCRENIDT 256
Db 180 VGVHVRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRENINA 239
QY 257 SHGDVVFAGDIEGSPAKDFALLTQCNTHTMTGTFGIWAAYLTGGDTIYLANYTLPDSP 316
Db 240 SRGDVVFAGNIEGSPAKDFALLTQCNTHTMTGTFGIWAAYLAGDGTIYLANYTLPDSP 299
QY 317 FLKIFKPEAAFLPEWTGIAADLSPLLK 343
Db 300 FLKVFPEAAFLPEWVGIPADLSPLLK 326
RESULT 8
AAE16623
ID AAE16623 standard; Protein; 353 AA.
XX
XX AAE16623;
XX

```



QY 42 PQVIPVLAST-----SKALGPSQLRGWMTINAIQRLGNQMGVEYATLYALAKMNGRPA 93  
 Db 51 PVALFCLAGTPVHPNADSDCPKHPASFGTWTIYPDGRFGNQMGQYATLLAALNGRQA 110  
 QY 94 FIPAQMSTLAPIFRITLPVLHLSATASRIQWNYHLNDWMEEEYRHIPPQGYVRFYGYPC 153  
 Db 111 FIOPAMHVLAPVFRITLPVLAPEDVRHAPWRELEHDWSEDYAHLKE-PWLKLTGFPC 169  
 QY 154 SWTFYHHLRQELIQLQFTHLHDHVRERAKFLRGLOV--NGSRPGTFVGVHVRGDDYVHVP 211  
 Db 170 SWTFHHLRQELIQLQFTHLHDHVRERAKFLRGLOV--NGSRPGTFVGVHVRGDDYVHVP 229  
 QY 212 KWKGVVADRRYLQOALDWFRRARYSSLIFFVVTSGMAWCRENIDTSHGDDYVAGDGIEGS 271  
 Db 230 KWKGVVGDGRYLQOAMDWRFRARYEAPFVVTSGMEWCNRKIDTSGDYVIFAGDGRQA 289  
 QY 272 PAKDFALLTCNHTIMTIGTFGIWAAYLTGSDTIYLANFTLPDSPFLKIKPPEAAFLPEW 331  
 Db 332 TGTAAADLSPL 341  
 Db 350 VGINADLSPL 359

RESULT 10  
 AAW30630  
 ID AAW30630 standard; Protein; 365 AA.  
 AC AAW30630;  
 AC XX  
 DT 01-APR-1999 (first entry)  
 XX  
 DE Swine alpha-1,2-fucosyltransferase 1 protein.  
 XX  
 KW Swine; pig; alpha-1,2-fucosyltransferase 1; FUT1; resistance;  
 KW Escherichia coli; infection; oedema; postweaning diarrhoea;  
 KW intestinal disorder; polymorphism.  
 XX  
 OS Sus scrofa.  
 XX  
 PN W09853102-A1.  
 XX  
 PD 26-NOV-1998.  
 XX  
 PF 20-MAY-1998; 98WO-US10318.  
 XX  
 PR 20-MAY-1997; 97US-0047181.  
 XX  
 PA (BIOT-) BIOTECHNOLOGY RES & DEV CORP.  
 PA (SWIT-) SWISS FEDERAL INST TECHNOLOGY ZURICH.  
 PA (USDA ) US SEC OF AGRIC.  
 XX  
 PI Bosworth BT, Vogeli P;  
 XX  
 DR WPI: 1999-059746/05.  
 DR N-PSDB; AAX03811.  
 XX  
 PT New method of identifying swine that are resistant to intestinal  
 PT colonisation by Escherichia coli - comprises use of genetic  
 PT polymorphic markers, used for breeding swine resistant to  
 PT Escherichia coli-related diseases  
 XX  
 PS Example 6; Fig 1; 35pp; English.  
 XX  
 CC A method has been developed for the identification of swine that are  
 CC resistant to intestinal colonisation by E. coli. The method comprises  
 CC determining whether a genetic polymorphism associated with resistance to  
 CC colonisation is present in a swine sample, and then inferring that the  
 CC swine is resistant if it is homozygous for the polymorphism. The method  
 CC uses the swine alpha-1,2-fucosyltransferase (FUT1) polymorphism. The  
 CC present sequence represents swine FUT1. The method enables the breeding

CC of swine that are resistant to E. coli-related diseases. This method  
 CC comprises breeding swine that have a genetic polymorphism in the FUT1  
 CC gene. More particularly, the identification method identifies swine that  
 CC are resistant to E. coli-related intestinal disorders if, in a sample  
 CC taken, the only nitrogen base at residue 307 in the FUT1 gene is adenine.  
 CC Larger amplified fragments from the assay can be used for RFLP analysis,  
 CC and the assay itself is used as a basis for a kit, applied to swine of  
 CC any age, in detecting polymorphisms associated with E. coli F18  
 CC receptors. The polymorphisms are useful in developing drugs to treat  
 CC swine with E. coli-related diseases. However, a mutated form of the  
 CC porcine FUT1 gene may interfere with the normal enzyme and prevent it  
 CC from producing the intestinal receptor for F18. The detection of  
 CC polymorphic markers in the method disclosed enables the detection and  
 CC treatment of E. coli-related intestinal diseases in swine, where there  
 CC has been no success using antibiotics due to unsuccessful prophylaxis.  
 XX

SO Sequence 365 AA;

Query Match 58.5%; Score 1081.5; DB 20; Length 365;  
 Best Local Similarity 65.5%; Pred. No. 3.3e-105;

Matches 203; Conservative 30; Mismatches 66; Indels 11; Gaps 3;

QY 42 PQVIPVLAST-----SKALGPSQLRGWMTINAIQRLGNQMGVEYATLYALAKMNGRPA 93  
 Db 51 PVALFCLAGTPVHPNADSDCPKHPASFGTWTIYPDGRFGNQMGQYATLLAALNGRQA 110  
 QY 94 FIPAQMSTLAPIFRITLPVLHLSATASRIQWNYHLNDWMEEEYRHIPPQGYVRFYGYPC 153  
 Db 111 FIOPAMHVLAPVFRITLPVLAPEDVRHAPWRELEHDWSEDYAHLKE-PWLKLTGFPC 169  
 QY 154 SWTFYHHLRQELIQLQFTHLHDHVRERAKFLRGLOV--NGSRPGTFVGVHVRGDDYVHVP 211  
 Db 170 SWTFHHLRQELIQLQFTHLHDHVRERAKFLRGLOV--NGSRPGTFVGVHVRGDDYVHVP 229  
 QY 212 KWKGVVADRRYLQOALDWFRRARYSSLIFFVVTSGMAWCRENIDTSHGDDYVAGDGIEGS 271  
 Db 230 KWKGVVGDGRYLQOAMDWRFRARYEAPFVVTSGMEWCNRKIDTSGDYVIFAGDGRQA 289  
 QY 272 PAKDFALLTCNHTIMTIGTFGIWAAYLTGSDTIYLANFTLPDSPFLKIKPPEAAFLPEW 331  
 Db 290 PARDFALLVCNHTIMTIGTFGIWAAYLAGSDTIYLANFTLFTSSFLKIKPPEAAFLPEW 349  
 QY 332 TGTAAADLSPL 341  
 Db 350 VGINADLSPL 359

RESULT 11  
 AAR13751

ID AAR13751 standard; Protein; 365 AA.

AC AAR13751;

DT 07-NOV-1991 (first entry)

DE GDP-Fuc:beta-D-galactoside alpha(1,2)-fucosyltransferase.

KW Glycosyltransferase.

OS Homo sapiens.

PN W09112340-A.

PD 22-AUG-1991.

PF 14-FEB-1991; 91WO-US00899.

PR 12-DEC-1990; 90US-0627621.

PR 14-FEB-1990; 90US-0479658.

PR 14-FEB-1990; 90US-0480133.

PA (UNMI ) UNIV OF MICHIGAN.

XX





```

OS Homo sapiens.
XX W09524495-A1.
XX 14-SEP-1995.
XX 24-JAN-1995; 95WO-US00967.
XX 09-MAR-1994; 94US-0208889.
XX (ABBO ) ABBOTT LAB.
XX Cummings RD, Kopchik JJ, Moremen KW, Mukerji P;
XX Pierce JM, Prieto PA, Smith DF;
XX WPI; 1995-328284/42.
XX N-PSDB; AAT01083.
XX New transgenic non-human mammal milk prods - contg. heterologous
XX components produced as secondary gene prods. of an heterologous gene
XX Example 1; Page 62-64; 83pp; English.
XX 2-Alpha-fucosyltransferase (AAR70422) is encoded by a cDNA clone
XX (AAT01083) isolated from a human epidermal carcinoma A431 library.
XX The enzyme can be expressed in the milk of a transgenic mammal,
XX esp. cow. This allows large-scale prodn. of oligosaccharides and
XX glycosylated proteins and lipids in the milk.
XX Sequence 365 AA;

Query Match 58.4%; Score 1080.5; DB 16; Length 365;
Best Local Similarity 59.7%; Pred. No. 4.2e-105;
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;

QY 12 MAHPILFVFTVSTIFVHQQR-----LAKIQAMWELPQIPVL-----ASTSK 53
DB 11 LAFLVCVLSVIFFLHIHQDSFPHGLGSLICPDRLVTPVPAIFCLPGTAMGNASSC 70
QY 54 ALGPSQLRGMTTINAGRLGNOMGEYATLALAKNGRPAFIPAQMHSTLAPIFRITLPV 113
DB 71 PQHPASLSGTWTVYPNGRFGNQMGQYATLLAQLNGRRRAFILPAMHAALAPVFRITLPV 130
QY 114 LHSATASRIPWNYHLNDWMEEEYRHPPGVEYVFTGPGCSWTFFYHHLRQILQELFDLHD 173
DB 131 LAPEVDSRTFWRELQLDHWMSEYADL-RDFFLKLSGFPSCSWTFHHLRQIRREFTLHD 189
QY 174 HVREAAQKFLRGLQV--NGSRPGTFVGVHVRGDDYVHVMKGVVADRRYLOQALDWF 231
DB 190 HLREAAQSVLGQLRLGRTGDRPRTFVGVHVRGDDYLOVMPQKGVVGDSDAYLRQAMDWF 249
QY 232 RARYSSLIFFVTSNGMAWCARENIDTSHGDDVVFAGDGTGEGSPAKDFALLTCQNHITMIGT 291
DB 250 RARAEAPFVVVTSNGMEWCARENIDTSQDVTFFAGDGEATPWKDFALLTCQNHITMIGT 309
QY 292 FGIWAAYLTGGDTIYLANYLTPDSPFLKIFKPEAFLEPWTGIAADLSPL 341
DB 310 FGFWAAYLAGGDTIYLANFLIPDSEFLKIFKPEAFLEPWTGIAADLSPL 359

RESULT 14
AAR70421
ID AAR70421 standard; Protein; 365 AA.
XX AAR70421;
XX 27-FEB-1996 (first entry)
XX 2-Alpha-fucosyltransferase.
XX GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyltransferase;
XX glycosyltransferase; oligosaccharide; glycoprotein; glycolipid;
XX transgenic animal; cattle; glycosylation; milk.

```

```

XX Homo sapiens.
XX W09524488-A1.
XX 14-SEP-1995.
XX 24-JAN-1995; 95WO-US01147.
XX 09-MAR-1994; 94US-0209132.
XX (ABBO ) ABBOTT LAB.
XX Cummings RD, Kopchik JJ, Moremen KW, Mukerji P;
XX Pierce JM, Prieto PA, Smith DF;
XX WPI; 1995-328279/42.
XX N-PSDB; AAT01082.
XX Transgenic animal expressing heterologous catalyst - used in
XX metabolic prodn esp. glycosyltransferase for prodn. of
XX oligosaccharide(s) and glyco-conjugate(s) in the milk, useful in
XX foods, pharmaceuticals, etc.
XX Example 1; Page 62-64; 84pp; English.
XX Alpha-1,2-fucosyltransferase (AAR70421) is encoded by a cDNA clone
XX (AAT01082) isolated from a human epidermal carcinoma cell line A431
XX cDNA library. The enzyme can be expressed in the milk of a
XX transgenic mammal, esp. cow. This allows large-scale prodn. of
XX oligosaccharides or glycosylated proteins and lipids in the milk.
XX Sequence 365 AA;

Query Match 58.4%; Score 1080.5; DB 16; Length 365;
Best Local Similarity 59.7%; Pred. No. 4.2e-105;
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;

QY 12 MAHPILFVFTVSTIFVHQQR-----LAKIQAMWELPQIPVL-----ASTSK 53
DB 11 LAFLVCVLSVIFFLHIHQDSFPHGLGSLICPDRLVTPVPAIFCLPGTAMGNASSC 70
QY 54 ALGPSQLRGMTTINAGRLGNOMGEYATLALAKNGRPAFIPAQMHSTLAPIFRITLPV 113
DB 71 PQHPASLSGTWTVYPNGRFGNQMGQYATLLAQLNGRRRAFILPAMHAALAPVFRITLPV 130
QY 114 LHSATASRIPWNYHLNDWMEEEYRHPPGVEYVFTGPGCSWTFFYHHLRQILQELFDLHD 173
DB 131 LAPEVDSRTFWRELQLDHWMSEYADL-RDFFLKLSGFPSCSWTFHHLRQIRREFTLHD 189
QY 174 HVREAAQKFLRGLQV--NGSRPGTFVGVHVRGDDYVHVMKGVVADRRYLOQALDWF 231
DB 190 HLREAAQSVLGQLRLGRTGDRPRTFVGVHVRGDDYLOVMPQKGVVGDSDAYLRQAMDWF 249
QY 232 RARYSSLIFFVTSNGMAWCARENIDTSHGDDVVFAGDGTGEGSPAKDFALLTCQNHITMIGT 291
DB 250 RARAEAPFVVVTSNGMEWCARENIDTSQDVTFFAGDGEATPWKDFALLTCQNHITMIGT 309
QY 292 FGIWAAYLTGGDTIYLANYLTPDSPFLKIFKPEAFLEPWTGIAADLSPL 341
DB 310 FGFWAAYLAGGDTIYLANFLIPDSEFLKIFKPEAFLEPWTGIAADLSPL 359

RESULT 15
AAR90572
ID AAR90572 standard; Protein; 365 AA.
XX AAR90572;
XX 08-APR-1996 (first entry)
XX Human H-transferase.

```



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:03:22 : Search time 14.1436 Seconds  
(without alignments)  
715.625 Million cell updates/sec

Title: US-10-040-863-11  
Perfect score: 1850  
Sequence: 1 MLVQMPFSPFMAHFLFV.....AAFLPWTGIAADLSPLK 344

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1850	100.0	344	4	US-09-298-886-11
2	1832.5	99.1	343	1	US-08-395-800A-8
3	1832.5	99.1	343	4	US-09-254-077A-8
4	1528	82.6	340	4	US-09-254-077A-6
5	1528	82.6	340	4	US-09-254-077A-7
6	1509.5	81.6	347	4	US-09-254-077A-9
7	1475	79.7	380	4	US-09-298-886-8
8	1376	74.4	353	4	US-09-298-886-10
9	1093.5	59.1	373	4	US-09-254-077A-12
10	1080.5	58.4	365	1	US-07-914-281-6
11	1080.5	58.4	365	1	US-08-393-246-6
12	1080.5	58.4	365	1	US-08-273-411-1
13	1080.5	58.4	365	1	US-08-535-058A-6
14	1080.5	58.4	365	1	US-08-395-800A-6
15	1080.5	58.4	365	1	US-08-395-800A-10
16	1080.5	58.4	365	2	US-08-696-731-6
17	1080.5	58.4	365	4	US-09-042-531-6
18	1080.5	58.4	365	4	US-09-042-077A-11
19	1080.5	58.4	365	5	PCT-US91-00899-12
20	1076.5	58.2	333	5	PCT-US91-00899-11
21	1076.5	58.1	365	4	US-09-151-592-2
22	1075.5	58.1	365	4	US-09-254-077A-10
23	809.5	43.8	222	1	US-08-395-800A-2
24	420	22.7	102	1	US-08-395-800A-3
25	144.5	7.8	300	4	US-09-433-598-2
26	100	5.4	355	4	US-09-390-131-5
27	92	5.0	575	3	US-08-913-805A-2

28	92	5.0	575	4	US-09-442-629-2	Sequence 2, Appli
29	90	4.9	575	3	US-08-913-805A-10	Sequence 10, Appl
30	90	4.9	575	4	US-09-442-629-10	Sequence 10, Appl
31	87.5	4.7	365	3	US-08-978-741-2	Sequence 2, Appli
32	87.5	4.7	365	4	US-09-333-729A-3	Sequence 2, Appli
33	87.5	4.7	397	3	US-08-978-741-6	Sequence 6, Appli
34	87.5	4.7	397	4	US-09-333-729A-7	Sequence 7, Appli
35	82.5	4.5	183	1	US-08-167-035-33	Sequence 33, Appl
36	82.5	4.5	183	2	US-08-208-887A-33	Sequence 33, Appl
37	82.5	4.5	183	2	US-08-539-005-33	Sequence 33, Appl
38	82.5	4.5	183	4	US-09-280-598-35	Sequence 35, Appl
39	82.5	4.5	1048	4	US-08-887-534A-85	Sequence 85, Appl
40	81.5	4.4	607	1	US-07-959-943-7	Sequence 7, Appli
41	81.5	4.4	630	1	US-07-959-943-9	Sequence 9, Appli
42	81.5	4.4	653	1	US-07-782-298-2	Sequence 2, Appli
43	80	4.3	604	2	US-08-735-041A-4	Sequence 4, Appli
44	80	4.3	604	3	US-09-190-476B-4	Sequence 4, Appli
45	80	4.3	604	3	US-09-190-889A-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1  
US-09-298-886-11  
; Sequence 11, Application US/09298886  
; Patent No. 5329170  
; GENERAL INFORMATION:  
; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: 8511-029  
; CURRENT APPLICATION NUMBER: US/09/298,886  
; CURRENT FILING DATE: 1999-04-26  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 344  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-298-886-11

Query Match 100.0%; Score 1850; DB 4; Length 344;  
Best Local Similarity 100.0%; Pred. No. 4.9e-204;  
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLVQMPFSPFMAHFLFVFTVSTIFHVQORLAKIQAMHQLVQIPVLASTSKALGPSOL	60
DB	1	MLVQMPFSPFMAHFLFVFTVSTIFHVQORLAKIQAMHQLVQIPVLASTSKALGPSOL	60
QY	61	RGMTINAIGRGNOMGEYATLYALAKMNGREAFIPQAQMHTLAPIFRITLPLVLSATAS	120
DB	61	RGMTINAIGRGNOMGEYATLYALAKMNGREAFIPQAQMHTLAPIFRITLPLVLSATAS	120
QY	121	RIPWQNYHLNDWMEYRHIPPEGVRTGYPCSWTFYHHLRQETLQETLHDHVRERQAQ	180
DB	121	RIPWQNYHLNDWMEYRHIPPEGVRTGYPCSWTFYHHLRQETLQETLHDHVRERQAQ	180
QY	181	KFLRLQVNGSRPGTFVGVVHRRGDYVHVMPKVGKVGVVADREYLOQALDWFARVSSLIIF	240
DB	181	KFLRLQVNGSRPGTFVGVVHRRGDYVHVMPKVGKVGVVADREYLOQALDWFARVSSLIIF	240
QY	241	VVTSNGMAWCRENIDTSHGDVVVAGDGIEGSPAKDFALLITQCNHITMITIGTFGIWAAYLT	300
DB	241	VVTSNGMAWCRENIDTSHGDVVVAGDGIEGSPAKDFALLITQCNHITMITIGTFGIWAAYLT	300
QY	301	GGDTIYLANIYLPSPFLKIFKPEAAFLPEWTGIAADLSPLK 344	
DB	301	GGDTIYLANIYLPSPFLKIFKPEAAFLPEWTGIAADLSPLK 344	

RESULT 2

US-08-395-800A-8

; Sequence 8, Application US/08395800A  
; Patent No. 5807732

; GENERAL INFORMATION:

; APPLICANT: LOWE, JOHN B

; APPLICANT: LENNON, GREGORY

; APPLICANT: ROQUIER, SYLVIE

; APPLICANT: GIORGI, DOMINIQUE

; APPLICANT: KELLY, ROBERT J

; TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE

; TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES/DNA SEQUENCES ENCODING THE

; TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF

; TITLE OF INVENTION: GENOTYPING A PERSON

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSER: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT

; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400

; CITY: ARLINGTON

; STATE: VIRGINIA

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/395,800A

; FILING DATE: 28-FEB-1995

; CLASSIFICATION: 435

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 413-3000

; TELEFAX: (703) 413-2220

; TELEX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 343 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-395-800A-8

Query Match 99.1%; Score 1832.5; DB 1; Length 343;  
Best Local Similarity 99.7%; Pred. No. 5e-202;  
Matches 343; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MLVQMPFSFPMAHFILFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSOL 60  
|||||

Db 1 MLVQMPFSFPMAHFILFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSOL 60  
|||||

QY 61 RGMWTINAIGRLGNOMGEYATLYALAKMNGRPAFIPAQMHSHTLAFIRITLPLVLSATAS 120  
|||||

Db 61 RGMWTINAIGRLGNOMGEYATLYALAKMNGRPAFIPAQMHSHTLAFIRITLPLVLSATAS 120  
|||||

QY 121 RIPWQNYHLNDWMEEEYRHIPEGEYVRFYPCSWTFYHHLRQELQETLHDHVREBAQ 180  
|||||

Db 121 RIPWQNYHLNDWMEEEYRHIPEGEYVRFYPCSWTFYHHLRQELQETLHDHVREBAQ 179  
|||||

QY 181 KFLRGLQVNGSRPGTFVGVHVRGQYVHVMPKVGKVADRRYLOQALDWFARVSSLIIF 240  
|||||

Db 180 KFLRGLQVNGSRPGTFVGVHVRGQYVHVMPKVGKVADRRYLOQALDWFARVSSLIIF 239  
|||||

QY 241 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNHNTIMTIGTFCGIAWAAYLT 300  
|||||

Db 240 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNHNTIMTIGTFCGIAWAAYLT 299  
|||||

QY 301 GGDITLYLANYILPDSFPLKIEKPEAAFLPEWTGTAAADLSPLKH 344  
|||||

Db 300 GGDITLYLANYILPDSFPLKIEKPEAAFLPEWTGTAAADLSPLKH 343  
|||||

RESULT 3

US-09-254-077A-8

; Sequence 8, Application US/09254077A

; Patent No. 6399758

; GENERAL INFORMATION:

; APPLICANT: SANDRIN, MAURO S.

; APPLICANT: MCKENZIE, IAN C. F.

; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES

; FILE REFERENCE: 30562.5USWO

; CURRENT APPLICATION NUMBER: US/09/254.077A

; CURRENT FILING DATE: 1999-06-11

; PRIOR APPLICATION NUMBER: PCT/AU97/00540

; PRIOR FILING DATE: 1997-08-22

; PRIOR APPLICATION NUMBER: PO 1823

; PRIOR FILING DATE: 1996-08-23

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 8

; LENGTH: 343

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-254-077A-8

Query Match 99.1%; Score 1832.5; DB 4; Length 343;  
Best Local Similarity 99.7%; Pred. No. 5e-202;  
Matches 343; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MLVQMPFSFPMAHFILFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSOL 60  
|||||

Db 1 MLVQMPFSFPMAHFILFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSOL 60  
|||||

QY 61 RGMWTINAIGRLGNOMGEYATLYALAKMNGRPAFIPAQMHSHTLAFIRITLPLVLSATAS 120  
|||||

Db 61 RGMWTINAIGRLGNOMGEYATLYALAKMNGRPAFIPAQMHSHTLAFIRITLPLVLSATAS 120  
|||||

QY 121 RIPWQNYHLNDWMEEEYRHIPEGEYVRFYPCSWTFYHHLRQELQETLHDHVREBAQ 180  
|||||

Db 121 RIPWQNYHLNDWMEEEYRHIPEGEYVRFYPCSWTFYHHLRQELQETLHDHVREBAQ 179  
|||||

QY 181 KFLRGLQVNGSRPGTFVGVHVRGQYVHVMPKVGKVADRRYLOQALDWFARVSSLIIF 240  
|||||

Db 180 KFLRGLQVNGSRPGTFVGVHVRGQYVHVMPKVGKVADRRYLOQALDWFARVSSLIIF 239  
|||||

QY 241 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNHNTIMTIGTFCGIAWAAYLT 300  
|||||

Db 240 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNHNTIMTIGTFCGIAWAAYLT 299  
|||||

QY 301 GGDITLYLANYILPDSFPLKIEKPEAAFLPEWTGTAAADLSPLKH 344  
|||||

Db 300 GGDITLYLANYILPDSFPLKIEKPEAAFLPEWTGTAAADLSPLKH 343  
|||||

RESULT 4

US-09-254-077A-6

; Sequence 6, Application US/09254077A

; Patent No. 6399758

; GENERAL INFORMATION:

; APPLICANT: SANDRIN, MAURO S.

; APPLICANT: MCKENZIE, IAN C. F.

; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES

; FILE REFERENCE: 30562.5USWO

; CURRENT APPLICATION NUMBER: US/09/254.077A

; CURRENT FILING DATE: 1999-06-11

; PRIOR APPLICATION NUMBER: PCT/AU97/00540

; PRIOR FILING DATE: 1997-08-22

; PRIOR APPLICATION NUMBER: PO 1823

; PRIOR FILING DATE: 1996-08-23

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 340

; TYPE: PRT

; ORGANISM: Sus scrofa

US-09-254-077A-6

```
Query Match      82.6%; Score 1528; DB 4; Length 340;
Best Local Similarity 82.0%; Pred. No. 5e-167;
Matches 282; Conservative 24; Mismatches 34; Indels 4; Gaps 2;

QY 1 MLVQMPFSPMAHFLVFVFTSTIFHVQORLAKIQAMWELPQVQIPVLASTSKALGPSOL 60
DB 1 MLSMQASFFFTGPFILFVFTASTIFHLQORVWKIQPTWELQM--VTQVTTESPSPQL 57

QY 61 RGMWTINAIGRLGNQMGEXATYALAKMNGRPAFIPAQMHSHTLAPIFRITLPLVLSHATAS 120
DB 58 KGMWTINAIGRLGNQMGEXATYALAKMNGRPAFIPPEMHSHTLAPIFRITLPLVLSHASTAR 117

QY 121 RIPQWYHLNDWMEERYRHLPPEYVFTGYPGCSWTFYHHLRQEILOEFTLHDHVEEAQ 180
DB 118 RIPQWYHLNDWMEERYRHLPPEYVFTGYPGCSWTFYHHLRQEILOEFTLHDHVEEAQ 176

QY 181 KFLRGVNSRGRTFVGVHVRGDDYVHVMKPKVKGWVADRRYLOQALDWFRRARYSSLPF 240
DB 177 DFLRGLRVNSRGRTFVGVHVRGDDYVHVMKPKVKGWVADRRYLOQALDWFRRARYSPVF 236

QY 241 VVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFCIWAAYLT 300
DB 237 VVTSNGMAWCRENINASRGDVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFCIWAAYLA 296

QY 301 GGTIYLANVTLDPSPFLKIFKPEAAFLPEWTGIAADLSPLLKH 344
DB 297 GGTIYLANVTLDPSPFLKIFKPEAAFLPEWIGIEADLSPLLKH 340

RESULT 5
US-09-254-077A-7
; Sequence 7, Application US/09254077A
; Patent No. 6399758
; GENERAL INFORMATION:
; APPLICANT: SANDRIN, MAURO S.
; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES
; FILE REFERENCE: 30562.5USWO
; CURRENT APPLICATION NUMBER: US/09/254,077A
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: PCT/AU97/00540
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: PO 1823
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-254-077A-7

Query Match      82.6%; Score 1528; DB 4; Length 340;
Best Local Similarity 82.0%; Pred. No. 5e-167;
Matches 282; Conservative 24; Mismatches 34; Indels 4; Gaps 2;

QY 1 MLVQMPFSPMAHFLVFVFTSTIFHVQORLAKIQAMWELPQVQIPVLASTSKALGPSOL 60
DB 1 MLSMQASFFFTGPFILFVFTASTIFHLQORVWKIQPTWELQM--VTQVTTESPSPQL 57

QY 61 RGMWTINAIGRLGNQMGEXATYALAKMNGRPAFIPAQMHSHTLAPIFRITLPLVLSHATAS 120
DB 58 KGMWTINAIGRLGNQMGEXATYALAKMNGRPAFIPPEMHSHTLAPIFRITLPLVLSHASTAR 117

QY 121 RIPQWYHLNDWMEERYRHLPPEYVFTGYPGCSWTFYHHLRQEILOEFTLHDHVEEAQ 180
DB 118 RIPQWYHLNDWMEERYRHLPPEYVFTGYPGCSWTFYHHLRQEILOEFTLHDHVEEAQ 176

QY 181 KFLRGVNSRGRTFVGVHVRGDDYVHVMKPKVKGWVADRRYLOQALDWFRRARYSSLPF 240
DB 177 DFLRGLRVNSRGRTFVGVHVRGDDYVHVMKPKVKGWVADRRYLOQALDWFRRARYSPVF 236

QY 241 VVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFCIWAAYLT 300

RESULT 6
US-09-254-077A-9
; Sequence 9, Application US/09254077A
; Patent No. 6399758
; GENERAL INFORMATION:
; APPLICANT: SANDRIN, MAURO S.
; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES
; FILE REFERENCE: 30562.5USWO
; CURRENT APPLICATION NUMBER: US/09/254,077A
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: PCT/AU97/00540
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: PO 1823
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Lepus Sp.
US-09-254-077A-9

Query Match      81.6%; Score 1509.5; DB 4; Length 347;
Best Local Similarity 81.3%; Pred. No. 6.9e-165;
Matches 283; Conservative 18; Mismatches 42; Indels 5; Gaps 3;

QY 1 MLVQMPFSPMAHFLVFVFTSTIFHVQORLAKIQAMWELPQVQIPVLASTSKALG 56
DB 1 MSTAQVFPFAPFVHVILFVFTASTIFHLQORVRIQPTWELLPAITPAVFRPTSORAP 60

QY 57 PSOLRGMTINAIGRLGNQMGEXATYALAKMNGRPAFIPAQMHSHTLAPIFRITLPLVLS 116
DB 61 SRPLGGMWTINAMGRGNQMGEXATYALAKMNGRPAFIPAQMHSHTLAPIFRISLPLVLS 120

QY 117 ATASRIPWQYHLNDWMEERYRHLPPEYVFTGYPGCSWTFYHHLRQEILOEFTLHDHVR 176
DB 121 STASRPVQWYHLNDWMEERYRHLPPEYVFTGYPGCSWTFYHHLRQEILOEFTLHDHVR 179

QY 177 EEAQKFLRGVNSRGRTFVGVHVRGDDYVHVMKPKVKGWVADRRYLOQALDWFRRARYS 236
DB 180 EEAQKFLRGVNSRGRTFVGVHVRGDDYVHVMKPKVKGWVADRRYLOQALDWFRRARYR 239

QY 237 SLIPVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFCIWA 296
DB 240 SPVFTVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFCIWA 299

QY 297 AYLGTGTIYLANVTLDPSPFLKIFKPEAAFLPEWTGIAADLSPLLKH 344
DB 300 AYLGTGTIYLANVTLDPSPFLKIFKPEAAFLPEWGINADLSPLLKH 347

RESULT 7
US-09-298-886-8
; Sequence 8, Application US/09298886
; Patent No. 6329170
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/298,886
; CURRENT FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
```



```

; TITLE OF INVENTION:  OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES:  14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET:  1755 Jefferson Davis Highway, Fourth Floor
; CITY:  Arlington
; STATE:  Virginia
; COUNTRY:  U.S.A.
; ZIP:  22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  PC-DOS/MS-DOS
; SOFTWARE:  PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/07/914,281
; FILING DATE:  19920720
; CLASSIFICATION:  530
; ATTORNEY/AGENT INFORMATION:
; NAME:  Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER:  31,451
; REFERENCE/DOCKET NUMBER:  2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:  (703)521-4500
; TELEFAX:  (703)486-2347
; TELEX:  248855 OPAT UR
; INFORMATION FOR SEQ ID NO:  6:
; SEQUENCE CHARACTERISTICS:
; LENGTH:  365 amino acids
; TYPE:  AMINO ACID
; TOPOLOGY:  unknown
; MOLECULE TYPE:  protein
; US-07-914-281-5

```

```

Query Match      58.4%; Score 1080.5; DB 1; Length 365;
Best Local Similarity 59.7%; Pred. No. 1.5e-115;
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;

QY  12 MAHFLFVETVSTIFHVQR-----LAKIQMWELPVOIPVL-----ASTSK 53
DB  11 LAFLLCVLSVFFLHGHDSFPHGLSLCPDRLVTPPVAIFCLPGTAMGNASSSC 70

QY  54 ALGPSOLRGMTINAGRLGNOMGEYATLYALAKMNGRPAFIPAQMNSTLAFIRITLPV 113
DB  71 PQHPASLSGTWTVYVNGRFGNQMGQVATLLAQLNGRRAFILPAMHAALAPVFRITLPV 130

QY  114 LHSATASRIPQWNYHLNDMEEYRHPGGEYVRFYGYPCSWTFYHHLRQEILOEFTLHD 173
DB  131 LAPEVDSRTPRELQLDHWMSEYADL-RDPFLKLSGFPSCSWTFHHLRQIRREFTLHD 189

QY  174 HVREEAQKFLRGLOV--NGSRPGTFVGVHVRGDIYQVMPKVKVGVADRRYLQQAALDWF 231
DB  190 HLREEAQSVLGQLRGTRGDRPRTFVGVHVRGDIYQVMPKVKVGVADRRYLQQAALDWF 249

QY  232 RARYSSLIFFVTSNGMAWCRENIDTSGHDVVFAGDIEGSPAKDFALLTCNHTIMTIGT 291
DB  250 RARHEAPVFTVTSNGMEWCENIDTSQGDVTFAGDQGEATPWKDFALLTCNHTIMTIGT 309

QY  292 FGIWAAYLTGGDTIYLANITLPDSPFLKFKPEAAFLPEWTGIAADLSPL 341
DB  310 FGFWAAYLAGGDTVYLANITLPDSPSEFLKFKPEAAFLPEWVGINADLSPL 359

```

```

RESULT 11
US-08-393-246-6
; Sequence 6, Application US/08393246
; Patent No. 559590
; GENERAL INFORMATION:
; APPLICANT:  LOWE, JOHN B.
; TITLE OF INVENTION:  METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; AND FOR THE ISOLATION
; OF GLYCOPOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION

```

```

; TITLE OF INVENTION:  OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES:  14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET:  1755 Jefferson Davis Highway, Fourth Floor
; CITY:  Arlington
; STATE:  Virginia
; COUNTRY:  U.S.A.
; ZIP:  22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  PC-DOS/MS-DOS
; SOFTWARE:  PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/08/393,246
; FILING DATE:
; CLASSIFICATION:  530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:  US 08/220,433
; FILING DATE:  30-MAR-1994
; APPLICATION NUMBER:  US 07/914,281
; FILING DATE:  20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME:  Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER:  31,451
; REFERENCE/DOCKET NUMBER:  2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:  (703)521-4500
; TELEFAX:  (703)486-2347
; TELEX:  248855 OPAT UR
; INFORMATION FOR SEQ ID NO:  6:
; SEQUENCE CHARACTERISTICS:
; LENGTH:  365 amino acids
; TYPE:  amino acid
; TOPOLOGY:  unknown
; MOLECULE TYPE:  protein
; US-08-393-246-6

```

```

Query Match      58.4%; Score 1080.5; DB 1; Length 365;
Best Local Similarity 59.7%; Pred. No. 1.5e-115;
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;

QY  12 MAHFLFVETVSTIFHVQR-----LAKIQMWELPVOIPVL-----ASTSK 53
DB  11 LAFLLCVLSVFFLHGHDSFPHGLSLCPDRLVTPPVAIFCLPGTAMGNASSSC 70

QY  54 ALGPSOLRGMTINAGRLGNOMGEYATLYALAKMNGRPAFIPAQMNSTLAFIRITLPV 113
DB  71 PQHPASLSGTWTVYVNGRFGNQMGQVATLLAQLNGRRAFILPAMHAALAPVFRITLPV 130

QY  114 LHSATASRIPQWNYHLNDMEEYRHPGGEYVRFYGYPCSWTFYHHLRQEILOEFTLHD 173
DB  131 LAPEVDSRTPRELQLDHWMSEYADL-RDPFLKLSGFPSCSWTFHHLRQIRREFTLHD 189

QY  174 HVREEAQKFLRGLOV--NGSRPGTFVGVHVRGDIYQVMPKVKVGVADRRYLQQAALDWF 231
DB  190 HLREEAQSVLGQLRGTRGDRPRTFVGVHVRGDIYQVMPKVKVGVADRRYLQQAALDWF 249

QY  232 RARYSSLIFFVTSNGMAWCRENIDTSGHDVVFAGDIEGSPAKDFALLTCNHTIMTIGT 291
DB  250 RARHEAPVFTVTSNGMEWCENIDTSQGDVTFAGDQGEATPWKDFALLTCNHTIMTIGT 309

QY  292 FGIWAAYLTGGDTIYLANITLPDSPFLKFKPEAAFLPEWTGIAADLSPL 341
DB  310 FGFWAAYLAGGDTVYLANITLPDSPSEFLKFKPEAAFLPEWVGINADLSPL 359

```

```

RESULT 12
US-08-273-411-1
; Sequence 1, Application US/08273411
; Patent No. 5625124

```

```

; GENERAL INFORMATION:
; APPLICANT: Falk, Per
; TITLE OF INVENTION: Animal Model for Gastro-Intestinal
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273.411
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 815-6508
; TELEFAX: (404) 815-6558
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..365
; OTHER INFORMATION: /note= "GDP-L-fucose:beta-D-Galactoside-2-alpha-L-fucosyl
; PUBLICATION INFORMATION:
; AUTHORS: Larsen, et al.
; JOURNAL: Proc. Nat'l Acad. Sci. USA
; VOLUME: 87
; PAGES: 6674-6678
; DATE: 1990
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 365
; US-08-273-411-1

```

```

Query Match 58.4%; Score 1080.5; DB 1; Length 365;
Best Local Similarity 59.7%; Pred. No. 1.5e-115;
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;

QY 12 MAHFILFVFTVSIFHFVQQR-----LAKIQMWELPVOIPVL-----ASTK 53
DB 11 LAFLLVCVLSVIFFLHIHQDSFPHGLSLCPDRRLVTPPPVAIFCLPGTAMGNPNASSC 70
QY 54 ALGPSQLRGMTWTINAIIGRLGNOMGEVATLYALAKMNGRAFIQAQMHSTLAFIFRITLPV 113
DB 71 PQHPASLSGTWVYPNGRGNOMGQVATLLAQLNGRAFIIPAMHAALAPVFRITLPV 130
QY 114 LHSATASRTIPWQNYHLNDWMEEFYRHPPEGVYRFTGYPGCSWTFYHHLRQETILQBFTHD 173
DB 131 LAPEVDSRTPWRRELQHDWMSSEYADL-RDPFLKLSGFPSCSWTFYHHLRQETIRREFTLHD 189
QY 174 HVREEAQKFLRGLOV--NGSRGFTGVGVHRRGDYVHVMPKVGVDYRRYLQOALDWF 231
DB 190 HUREBAQSVGLGRLGRTGDRPRTFVGVHRRGDYLVQVMPQKVGVDYSAVLRQAMDWF 249
QY 232 RARYSSLIFVVTISNGMAWCRENIDTSHGVDVVFAGDIEGSPAKDFALLTQCNTHTIMTGT 291

```

```

; APPLICANT: Lowe, John B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,058A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavallee, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-525-058A-6

```

```

Query Match 58.4%; Score 1080.5; DB 1; Length 365;
Best Local Similarity 59.7%; Pred. No. 1.5e-115;
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;

QY 12 MAHFILFVFTVSIFHFVQQR-----LAKIQMWELPVOIPVL-----ASTK 53
DB 11 LAFLLVCVLSVIFFLHIHQDSFPHGLSLCPDRRLVTPPPVAIFCLPGTAMGNPNASSC 70
QY 54 ALGPSQLRGMTWTINAIIGRLGNOMGEVATLYALAKMNGRAFIQAQMHSTLAFIFRITLPV 113
DB 71 PQHPASLSGTWVYPNGRGNOMGQVATLLAQLNGRAFIIPAMHAALAPVFRITLPV 130
QY 114 LHSATASRTIPWQNYHLNDWMEEFYRHPPEGVYRFTGYPGCSWTFYHHLRQETILQBFTHD 173
DB 131 LAPEVDSRTPWRRELQHDWMSSEYADL-RDPFLKLSGFPSCSWTFYHHLRQETIRREFTLHD 189
QY 174 HVREEAQKFLRGLOV--NGSRGFTGVGVHRRGDYVHVMPKVGVDYRRYLQOALDWF 231
DB 190 HUREBAQSVGLGRLGRTGDRPRTFVGVHRRGDYLVQVMPQKVGVDYSAVLRQAMDWF 249
QY 232 RARYSSLIFVVTISNGMAWCRENIDTSHGVDVVFAGDIEGSPAKDFALLTQCNTHTIMTGT 291

```



Db 250 RARHAPFVVVTSNGMEWCKENIDTSQDVTTFAGDGGQEAATPWKDFALLTCQNHMTIMTIGT 309  
QY 292 FGIWAAYLTGDDTYLYANVTLPDSPFLKIFKPEAAFLPEWTGIAADLSPL 341  
Db 310 FGIWAAYLAGDGTIVLANFTLPDSEFLKIFKPEAAFLPEWVGINADLSPL 359

## RESULT 14

US-08-395-800A-6  
: Sequence 6, Application US/08395800A  
: Patent No. 5807732  
: GENERAL INFORMATION:  
: APPLICANT: LOWE, JOHN B  
: APPLICANT: LENNON, GREGORY  
: APPLICANT: ROQUIER, SYLVIE  
: APPLICANT: GIORGI, DOMINIQUE  
: APPLICANT: KELLY, ROBERT J  
: TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE  
: TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE  
: TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF  
: TITLE OF INVENTION: GENOTYPING A PERSON  
: NUMBER OF SEQUENCES: 22  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT  
: STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
: CITY: ARLINGTON  
: STATE: VIRGINIA  
: COUNTRY: USA  
: ZIP: 22202  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent In Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/395,800A  
: FILING DATE: 28-FEB-1995  
: CLASSIFICATION: 435  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (703) 413-3000  
: TELEX: 248855 OPAT UR  
: INFORMATION FOR SEQ ID NO: 6:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 365 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
US-08-395-800A-6

Query Match 58.4%; Score 1080.5; DB 1; Length 365;  
Best Local Similarity 59.7%; Pred. No. 1.5e-115;  
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;  
QY 12 MAHFILFVFTVSTIFVHVOOR-----LAKIOAMWELPVQIPVL-----ASTSK 53  
Db 11 LAFLLCVLSVIFFLHLDHDSFPHGLGSLCPDRRLVTPPVAIFCLPGTAMGNASSC 70  
QY 54 ALGFSQLRGWMTNATIGRLGNQMGYATLYALAKMNGRPAPFAQMHSTLAPIRITLPV 113  
Db 71 PQHPASLSGTWTVYPNGREGNQGYATLLALAGLNGRRAPFILPAMHAALAPVFRITLPV 130  
QY 114 LHSATASRIPQNVHLNDWMEEEYRHPPGEYVRFYPCSWTFYHHLRQEIQLQEFTHD 173  
Db 131 LAPVDSTRTWRELQLDHDMSEEVADL-RDPFLKLSGFCPSWTFPHLRQIRREFTHD 189  
QY 174 HVREAAKFLRGLOV--NGSRPGTFVGVHVRGDDYVHVPKWKGVVADRYLQOALDNF 231  
Db 190 HLREAAQSVLQGLRLGTGDRPRTFVGVHVRGDDYLVQVQPKWGVVGDYSLRQAMDWF 249  
QY 232 RARYSSLIFFVVTSGMCAWCKENIDTSQDVTTFAGDGGQEAATPWKDFALLTCQNHMTIMTIGT 291  
Db 250 RARHAPFVVVTSNGMEWCKENIDTSQDVTTFAGDGGQEAATPWKDFALLTCQNHMTIMTIGT 309

QY 292 FGIWAAYLTGDDTYLYANVTLPDSPFLKIFKPEAAFLPEWTGIAADLSPL 341  
Db 310 FGIWAAYLAGDGTIVLANFTLPDSEFLKIFKPEAAFLPEWVGINADLSPL 359

## RESULT 15

US-08-395-800A-10  
: Sequence 10, Application US/08395800A  
: Patent No. 5807732  
: GENERAL INFORMATION:  
: APPLICANT: LOWE, JOHN B  
: APPLICANT: LENNON, GREGORY  
: APPLICANT: ROQUIER, SYLVIE  
: APPLICANT: GIORGI, DOMINIQUE  
: APPLICANT: KELLY, ROBERT J  
: TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE  
: TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE  
: TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF  
: TITLE OF INVENTION: GENOTYPING A PERSON  
: NUMBER OF SEQUENCES: 22  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT  
: STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
: CITY: ARLINGTON  
: STATE: VIRGINIA  
: COUNTRY: USA  
: ZIP: 22202  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent In Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/395,800A  
: FILING DATE: 28-FEB-1995  
: CLASSIFICATION: 435  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (703) 413-3000  
: TELEX: 248855 OPAT UR  
: INFORMATION FOR SEQ ID NO: 10:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 365 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
US-08-395-800A-10

Query Match 58.4%; Score 1080.5; DB 1; Length 365;  
Best Local Similarity 59.7%; Pred. No. 1.5e-115;  
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;  
QY 12 MAHFILFVFTVSTIFVHVOOR-----LAKIOAMWELPVQIPVL-----ASTSK 53  
Db 11 LAFLLCVLSVIFFLHLDHDSFPHGLGSLCPDRRLVTPPVAIFCLPGTAMGNASSC 70  
QY 54 ALGFSQLRGWMTNATIGRLGNQMGYATLYALAKMNGRPAPFAQMHSTLAPIRITLPV 113  
Db 71 PQHPASLSGTWTVYPNGREGNQGYATLLALAGLNGRRAPFILPAMHAALAPVFRITLPV 130  
QY 114 LHSATASRIPQNVHLNDWMEEEYRHPPGEYVRFYPCSWTFYHHLRQEIQLQEFTHD 173  
Db 131 LAPVDSTRTWRELQLDHDMSEEVADL-RDPFLKLSGFCPSWTFPHLRQIRREFTHD 189  
QY 174 HVREAAKFLRGLOV--NGSRPGTFVGVHVRGDDYVHVPKWKGVVADRYLQOALDNF 231  
Db 190 HLREAAQSVLQGLRLGTGDRPRTFVGVHVRGDDYLVQVQPKWGVVGDYSLRQAMDWF 249  
QY 232 RARYSSLIFFVVTSGMCAWCKENIDTSQDVTTFAGDGGQEAATPWKDFALLTCQNHMTIMTIGT 291  
Db 250 RARHAPFVVVTSNGMEWCKENIDTSQDVTTFAGDGGQEAATPWKDFALLTCQNHMTIMTIGT 309

QY 292 FGIWAAIYLTGGDTIYLANFTLPDSEFLKIKPEAAFLPEWTGIAADLSPL 341  
DB 310 FGFWAAIYLAGSDTYLANFTLPDSEFLKIKPEAAFLPEWVGINADLSPL 359

Search completed: May 27, 2003, 15:11:53  
Job time : 15.1436 secs

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:08:29 ; Search time 14.8627 Seconds  
(without alignments)  
2295.367 Million cell updates/sec

Title: US-10-040-863-11  
Perfect score: 1850  
Sequence: 1 MLVVQMPFSPMAHFLFV.....AAFLPWTGIAADLSPLKH 344

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications,AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

ALIGNMENTS

RESULT 1  
US-09-999-672-11  
; Sequence 11, Application US/09999672  
; Patent No. US20020127655A1  
; GENERAL INFORMATION:  
; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA-2 FUCCOSYLTRANSFERASE AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: 8511-029  
; CURRENT APPLICATION NUMBER: US/09/999,672  
; CURRENT FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: US/09/298,886  
; PRIOR FILING DATE: 1999-04-26  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 344  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-999-672-11

Query Match		100.0%;	Score 1850;	DB 10;	Length 344;
Best Local Similarity		100.0%;	Pred. No. 9,1e-183;		
Matches 344;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MLVVQMPFSPMAHFLFVFTVSTIFHVQORLAKIQAMWELPVOIPVLASTSKALGPSQL	60		
Db	1	MLVVQMPFSPMAHFLFVFTVSTIFHVQORLAKIQAMWELPVOIPVLASTSKALGPSQL	60		
QY	61	RGWMTINAGRLGNQMGYATLYALAKMNGRPAPFPAQMHSHTLAPIFRITLPVLHSATAS	120		
Db	61	RGWMTINAGRLGNQMGYATLYALAKMNGRPAPFPAQMHSHTLAPIFRITLPVLHSATAS	120		
QY	121	RIPQWNLNDWMEERHIPPGEYVRFYVFTGTPCSTWTFYHHLRQELQFTLHDHVRREAAQ	180		
Db	121	RIPQWNLNDWMEERHIPPGEYVRFYVFTGTPCSTWTFYHHLRQELQFTLHDHVRREAAQ	180		
QY	181	KFLRGLQVNGSRPCTFVGHVHVRGDIYHVMPKWKGVVADRRYQQALDWPFRARYSLIF	240		
Db	181	KFLRGLQVNGSRPCTFVGHVHVRGDIYHVMPKWKGVVADRRYQQALDWPFRARYSLIF	240		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1850	100.0	344	10	US-09-999-672-11
2	1850	100.0	344	12	US-10-040-863-11
3	1528	82.6	340	10	US-09-051-034A-2
4	1475	79.7	380	10	US-09-999-672-8
5	1475	79.7	380	12	US-10-040-863-8
6	1376	74.4	353	10	US-09-999-672-10
7	1376	74.4	353	12	US-10-040-863-10
8	1081.5	58.5	365	10	US-09-844-268-13
9	1081.5	58.5	365	9	US-09-844-705-13
10	1080.5	58.4	365	10	US-10-105-963-10
11	1080.5	58.4	365	10	US-09-863-475A-6
12	1075.5	58.1	365	10	US-09-051-034A-4
13	144.5	7.8	300	10	US-09-848-838-2
14	119.5	6.5	431	9	US-09-479-614-14
15	119.5	6.5	496	9	US-09-479-614-2
16	119.5	6.5	496	9	US-09-479-614-29
17	119.5	6.5	496	9	US-10-214-524-25
18	92	5.0	575	10	US-09-839-136-2
19	91	4.9	575	10	US-09-839-136-10

```

QY 241 VVTSNGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTIMTIGTGIWAAYLT 300
Db 241 VVTSNGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTIMTIGTGIWAAYLT 300
QY 301 GGDITILANTLSDSPFLKIFKPEAAFLPEWTGIAADLSPLKX 344
Db 301 GGDITILANTLSDSPFLKIFKPEAAFLPEWTGIAADLSPLKX 344

RESULT 2
US-10-040-863-11
; Sequence 11, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/298,886
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-040-863-11

```

```

Query Match 100.0%; Score 1850; DB 12; Length 344;
Best Local Similarity 100.0%; Pred. No. 9,1e-183;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MLVQMPFSPFMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSQL 60
Db 1 MLVQMPFSPFMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSQL 60
QY 61 RGMWTINAGRLGNQMGAYATLYALAKMNGRPAFIPQMHSTLAPIFRITLPLVLSHTAS 120
Db 61 RGMWTINAGRLGNQMGAYATLYALAKMNGRPAFIPQMHSTLAPIFRITLPLVLSHTAS 120
QY 121 RIPWQNYHLNDWMEERYRHIPPGYVFTGYPCSWTFYHHLRQELQEFTHDHYVREAAQ 180
Db 121 RIPWQNYHLNDWMEERYRHIPPGYVFTGYPCSWTFYHHLRQELQEFTHDHYVREAAQ 180
QY 181 KFLRGLOVNGSRPGTFVGVHVRGDYVHVMPKVKVGVVADRRYLOQALDWFRRYSSLI 240
Db 181 KFLRGLOVNGSRPGTFVGVHVRGDYVHVMPKVKVGVVADRRYLOQALDWFRRYSSLI 240
QY 241 VVTSNGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTIMTIGTGIWAAYLT 300
Db 241 VVTSNGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTIMTIGTGIWAAYLT 300
QY 301 GGDITILANTLSDSPFLKIFKPEAAFLPEWTGIAADLSPLKX 344
Db 301 GGDITILANTLSDSPFLKIFKPEAAFLPEWTGIAADLSPLKX 344

```

```

RESULT 3
US-09-051-034A-2
; Sequence 2, Application US/09051034A
; Patent No. US20010055584A1
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, IAN FARQUHAR CAMPBELL
; APPLICANT: SANDRIN, MAURO SERGIO
; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC
; TITLE OF INVENTION: GLYCOSYLTRANSFERASE
; FILE REFERENCE: 30562.6USWO
; CURRENT APPLICATION NUMBER: US/09/051,034A
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: PCT/AU97/00492

```

```

; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: 60/024,279
; PRIOR FILING DATE: 1996-08-21
; PRIOR APPLICATION NUMBER: P01402
; PRIOR FILING DATE: 1996-08-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Sus Domesticus
US-09-051-034A-2

```

```

Query Match 82.6%; Score 1528; DB 10; Length 340;
Best Local Similarity 82.0%; Pred. No. 1.7e-149;
Matches 282; Conservative 24; Mismatches 34; Indels 4; Gaps 2;

```

```

QY 1 MLVQMPFSPFMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSQL 60
Db 1 MLVQMPFSPFMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSQL 60
QY 61 RGMWTINAGRLGNQMGAYATLYALAKMNGRPAFIPQMHSTLAPIFRITLPLVLSHTAS 120
Db 61 RGMWTINAGRLGNQMGAYATLYALAKMNGRPAFIPQMHSTLAPIFRITLPLVLSHTAS 120
QY 121 RIPWQNYHLNDWMEERYRHIPPGYVFTGYPCSWTFYHHLRQELQEFTHDHYVREAAQ 180
Db 121 RIPWQNYHLNDWMEERYRHIPPGYVFTGYPCSWTFYHHLRQELQEFTHDHYVREAAQ 180
QY 181 KFLRGLOVNGSRPGTFVGVHVRGDYVHVMPKVKVGVVADRRYLOQALDWFRRYSSLI 240
Db 181 KFLRGLOVNGSRPGTFVGVHVRGDYVHVMPKVKVGVVADRRYLOQALDWFRRYSSLI 240
QY 241 VVTSNGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTIMTIGTGIWAAYLT 300
Db 241 VVTSNGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTIMTIGTGIWAAYLT 300
QY 301 GGDITILANTLSDSPFLKIFKPEAAFLPEWTGIAADLSPLKX 344
Db 301 GGDITILANTLSDSPFLKIFKPEAAFLPEWTGIAADLSPLKX 344

```

```

RESULT 4
US-09-999-672-8
; Sequence 8, Application US/09999672
; Patent No. US20020127655A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/999,672
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US/09/298,886
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-999-672-8

```

```

Query Match 79.7%; Score 1475; DB 10; Length 380;
Best Local Similarity 77.7%; Pred. No. 5.9e-144;
Matches 275; Conservative 30; Mismatches 37; Indels 12; Gaps 3;

```

```

QY 1 MLVQMPFSPFMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSQL 60
Db 1 MASQVPPFSPFLAHLFVFTVSTIHLQQRIVKQLPLSEKELPMTQMSSTNTSPEHAR 60
QY 50 STSKALGPSQLRGMWTINAGRLGNQMGAYATLYALAKMNGRPAFIPQMHSTLAPIFR 109

```

Db	61	RDSEQNGNELRGFMFTINSTIGRLNGOMGYSATLFAALRWNGELATIPASMHNALAPIERI	120
Qy	110	TLPLVLSATASRIPIWONYHLNDWMEDEYRHIIPPGYVFRGTGVCSTWTFVHHLRQILGEF	169
Db	121	SLPVLSDTAKKIPWQNYHLNDWMEERYRHI-PCGFVRSTGTGVCSTWTFVHHLRPEILKEF	179
Qy	170	TLHDHVRREAQKFLRGLGVNGSRPGFVGVHVRGRDGYVHVMKVKWGVVADRYLQQAALD	229
Db	180	TLHDHVRREAQKFLRGLGVNGSOPSTFGVHVRGRDGYVHVMNVKVGVDARYLEKALD	239
Qy	230	WFRARYSSLFVVTISNGMAWCNENIDTSGDGVVTFAGDGTGGSPAKDFALLTCQNTHTMTI	289
Db	240	MFRARYSSPFFVVTISNGMAWCNENINAKSDGVVTFAGDGTGGSPAKDFALLTCQNTHTMTI	299
Qy	290	GTFCINAAATLTGGDTIYLANYTLPDPSFFLKFKEPAAFLEPWETGTAADLSPLLK	343
Db	300	GTFCINAAATLTGGDTIYLANYTLPDPSFFLKFKEPAAFLEPWVGTPADLSPLLK	353

```

RESULT 5
US-10-040-863-8
; Sequence 8, Application US/10040863
; Patent NO. US2020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: CMI-SPECIFIC ALPHA-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 03/298,886
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 380
; TYPE: prt
; ORGANISM: Rattus norvegicus
US-10-040-863-8

```

```

RESULT 7
US-10-040-863-10
; Sequence 10, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: ERIC H. HOLMES et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 03/298,886
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-040-863-10

```

Query Match	74.4%;	Score 1376;	DB 12;	Length 353;
Best Local Similarity	78.9%;	Pred. No. 9e-134;		



```

;
; TELEPHONE: (703)521-4500
;
; TELEFAX: (703)456-2347
;
; TELE: 248855 OPAT UR
;
; INFORMATION FOR SEQ ID NO: 6:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 365 amino acids
;
; TYPE: amino acid
;
; TOPOLOGY: unknown
;
; MOLECULE TYPE: protein
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
;
US-09-863-475A-6

Query Match 58.4%; Score 1080.5; DB 10; Length 365;
Best Local Similarity 59.7%; Pred. No. 3.3e-103;
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps

QY 12 MAHFILFVFTVSIFVQQR-----LAKIQAMWELPVOIPVL-----ASTSK 53
Db 11 LAFLVCVLISVFLLHQDSFPHGLSLICDFRRLLVTPVAFICLPGTAMGPNASSC 71
QY 54 ALGPSOLRGMTTNAIGRLGNOMGEYATLYALAKMNGRPAFIPAQMHSHTLAPIFRITLPV 113
Db 71 PQHPASLSGWTVPVNGRCFGNQGVATLLAQLNGRRAFILPAMHAALAPVFRITLPV 130
QY 114 LHSATASRIPQWYHNLNDWMEBYRIIPGEYVFRFGYPCSWTFYHHLRQELIOETLHD 173
Db 131 LAPEVDSRTPEWRELQIHDNMSEYADL-RDPFKLSGFPCCSWTFHHLRQIRREFTLHD 189
QY 174 HVREEACKFLRLQV--NGSRPTEFVGHVRRGDYHVMPKVKWGVVADRRYLIQOALDNF 231
Db 190 HLREEAGSVLQURLGRTGDRPTFVGHVRRGDYLOVMPQRWGVVGDYSAYLRQAMDWF 249
QY 232 RARYSLIFVYTSNGMAWCRENIDTSHGDVVFAGDGIEGSPAKDFALLTCQNTIMTIGT 291
Db 250 RARHEAPVFTVTSNGHEWCENIDTSGDVTYFAGDGOEATPWKDFALLTCQNTIMTIGT 309
QY 292 FGTWAAYLTCGDITYILANYTLTPSPFLKIFKPEAAFLPWTGIAADLSPL 341
Db 310 FGFWAAYLAGGDTVYLANFTLPDSEFLKIFKPEAAFLPEWVGINADLSPL 359

RESULT 12
US-09-051-034A-4
; Sequence 4, Application US/09051034A
; Patent No. US20010055584A1
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, IAN FAROUHAR CAMPBELL
; APPLICANT: SANDRIN, MAURO SERGIO
; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC
; TITLE OF INVENTION: GLYCOSYLTRANSFERASE
; FILE REFERENCE: 30562.6USWO
; CURRENT APPLICATION NUMBER: US/09/051,034A
; CURRENT FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: PCT/AU97/00492
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: 60/024,279
; PRIOR FILING DATE: 1996-08-21
; PRIOR APPLICATION NUMBER: P01402
; PRIOR FILING DATE: 1996-08-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Sus Domesticus
US-09-051-034A-4

Query Match 58.1%; Score 1075.5; DB 10; Length 365;
Best Local Similarity 65.2%; Pred. No. 1.1e-102;
Matches 202; Conservative 30; Mismatches 67; Indels 1; Gaps

QY 42 PQVQIPVLAH-----SKALGPSQLRGWMT-NAIGRLGNOMGEYATLYALAKMGRPA 93

```

```
Db 51 PVAIFCLAGTVPVHNADSCPKHPASFSGTWITPOGRFGNQMGQYATLLAALNGRQA 110
Qy 94 FIPQMHSTLAPIRITLPLVLSATASRIPQNYHLNDWMEERYRHIPPGEYVRFYGYPC 153
Db 111 FLOPAMHVLAPVRITLPLVLAPEVDHAPRELELHDWSEDIYAHUKE-PWLKLTGPCC 169
Qy 154 SWTFYHHLRQELQEFTHLHDHVREAAQFLRGLOV--NGSRPGTFVGVHVRGDIYVHYMP 211
Db 170 SWTFPHHLRQELQEFTHLHDHVREAAQFLRGLOV--NGSRPGTFVGVHVRGDIYVHYMP 229
Qy 212 KWKGVVADRYLQALDWFARVSSLIFFVTSNGMACWRENIDTSGDVFAGDGBGS 271
Db 230 KWKGVVADRYLQALDWFARVSSLIFFVTSNGMACWRENIDTSGDVFAGDGBGS 289
Qy 272 PAKDFALLTQCNTHTMTTGTGFWAAAYLTGGDTIYLANFTLPSDFLKIPKPEAAFLPEW 331
Db 290 PARDFALLVQCNTHTMTTGTGFWAAAYLTGGDTIYLANFTLPSDFLKIPKPEAAFLPEW 349
Qy 332 TGIADLSPL 341
Db 350 VGINADLSPL 359

RESULT 13
US-09-848-838-2
; Sequence 2, Application US/09848838
; Patent No. US20020037570A1
; GENERAL INFORMATION:
; APPLICANT: Taylor, Diane
; APPLICANT: Wang, Ge
; APPLICANT: Paicic, Monica
; TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
; FILE REFERENCE: 07254-061002
; CURRENT APPLICATION NUMBER: US/09/848,838
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: US 09/433,598
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: US 60/107,268
; PRIOR FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-848-838-2
```

```
Query Match 7.8%; Score 144.5; DB 10; Length 300;
Best Local Similarity 22.9%; Pred. No. 1.3e-06;
Matches 61; Conservative 37; Mismatches 123; Indels 45; Gaps 8;

Qy 66 INAIGRLGNQMGYATLYALAKMNGRPAPFA-----QMHSTLAPIRITLPLVLSATAS 120
Db 6 VQICGGLGNQMGYATLYALAKMNGRPAPFA-----QMHSTLAPIRITLPLVLSATAS 120
Qy 121 RIPQNYHLNDWMEERYRHIPPGEYVRFYGYPCSWTFYHHLR 162
Db 66 IAKMQ--HLPKLVRLDALKYIGIDRVRSQEIYEPFKLLKPSRLTYFFCYQDPDYFDAIS 123
Qy 163 QEILQEFTHLHDHVREAAQFLRGLOV--NGSRPGTFVGVHVRGDIYVHYMP 213
Db 124 SLIKQTFPLPPPPNNKNNKEEYQKLS--LILAANKSVF--VHIRRGDYGVI----- 175
Qy 214 KWGVVADRYLQALDWFARVSSLIFFVTSNGMACWRENIDTSGDVFAGDGBGS 273
Db 176 --GCOLGIDYQKALEYAKVRPNMELEVFCEDIKFP--TQNLDLGYPFTDMITRKEEAY 232
Qy 274 KDFALLTQCNTHTMTTGTGFWAAAYLTGGDTIYLANFTLPSDFLKIPKPEAAFLPEW 349
Db 233 WDLMLMQCKHGILANSTYSWAAAYLTGGDTIYLANFTLPSDFLKIPKPEAAFLPEW 349

RESULT 14
US-09-848-838-2
```

```
US-09-479-614-14
; Sequence 14, Application US/09479614
; Publication No. US20030013183A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine
; APPLICANT: Weber, Eric
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods
; FILE REFERENCE: P-1047
; CURRENT APPLICATION NUMBER: US/09/479,614
; CURRENT FILING DATE: 2000-01-07
; EARLIER APPLICATION NUMBER: 60/115,033
; EARLIER FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Felis catus
US-09-479-614-14

Query Match 6.5%; Score 119.5; DB 9; Length 431;
Best Local Similarity 23.8%; Pred. No. 0.00084;
Matches 72; Conservative 27; Mismatches 111; Indels 93; Gaps 15;

Qy 47 VLASTSKALGPSQLRGMWMTINAIGRLGNQMGYATLYALAKMNGRPAPFAIPQMHSHT--- 102
Db 102 VSATMNFIPPTVKLFHSSCNPLDGTG-----STIQLCLLSG---YVPGDMETWLVLD 152
Qy 103 ---LAPIFRITLPLVLSATASRIPQNYHLNDWMEERYRHIPPGEYVRFYGYPCSWTFYH 159
Db 153 GOKATNIFPYTAP--GKQEGKVTSTHSELN-----ITQGEWVSQKTYTCQVTV-- 198
Qy 160 HLRQELQEFTHLHDHVREAAQFLRGLOVNGSRCTFTFVGVHVRGDIYVHYMPKWKGVVA 219
Db 199 -----QGTTFEDHARKCTESDPRGVSTYLSPPSPL-----DLYVHKSPKI--TCLV 243
Qy 220 DRYLQALDWFARVSSLIFFVTSNGM--AWCRNIDTSHGDVVFAGDGBGSFAKQFA 277
Db 244 D-----LANTDGMILTSRENGESVHPD-----PMV 269
Qy 278 LITOCNHTMTTGTGFWAAAYLTGGDTIYLANFTLPSD-----FLKI---FKPEA-A 326
Db 270 KKTQYNGTITVTSLPVDATDWEGET-YQCKVTHPLPKDIVRSIAKAPGRFPPEVYV 328
Qy 327 FLP 329
Db 329 FLP 331

RESULT 15
US-09-479-614-2
; Sequence 2, Application US/09479614
; Publication No. US20030013183A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine
; APPLICANT: Weber, Eric
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods
; FILE REFERENCE: P-1047
; CURRENT APPLICATION NUMBER: US/09/479,614
; CURRENT FILING DATE: 2000-01-07
; EARLIER APPLICATION NUMBER: 60/115,033
; EARLIER FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Felis catus
US-09-479-614-2

Query Match 6.5%; Score 119.5; DB 9; Length 496;
Best Local Similarity 23.8%; Pred. No. 0.001;
Matches 72; Conservative 27; Mismatches 111; Indels 93; Gaps 15;
```





GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:00:45 ; Search time 12.4655 Seconds  
(without alignments)  
2652.940 Million cell updates/sec

Title: US-10-040-863-11  
Perfect score: 1850  
Sequence: 1 MLVQMPFSPMAHFILEVF.....AAFLPEWTGIAADLSPLLK 344

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1832.5	99.1	343	2 A56098	alpha(1,2)fucosylt
2	1375.5	74.4	354	2 B56392	beta-galactoside a
3	1093.5	59.1	373	2 A56392	beta-galactoside a
4	1080.5	58.4	365	2 A36047	galactoside 2-alpha
5	623	33.7	159	2 S46424	galactoside 2-alpha
6	593	32.1	142	2 S51582	galactoside 2-alpha
7	195	10.5	281	2 T44328	hypothetical prote
8	181.5	9.8	363	2 T20745	hypothetical prote
9	176	9.5	443	2 H87911	protein B0205.4 [i
10	155	8.4	395	2 T32309	hypothetical prote
11	144.5	7.8	348	2 T31916	hypothetical prote
12	141.5	7.6	335	2 T32294	hypothetical prote
13	138	7.5	299	2 H71976	probable alpha(1,2
14	137.5	7.4	392	2 T25334	hypothetical prote
15	131.5	7.1	348	2 T02798	hypothetical prote
16	131	7.1	625	2 T21051	hypothetical prote
17	129.5	7.0	388	2 T32307	hypothetical prote
18	122	6.6	365	2 T32553	hypothetical prote
19	120.5	6.5	383	2 T20572	hypothetical prote
20	120	6.5	434	2 T26275	hypothetical prote
21	118.5	6.4	353	2 T25390	hypothetical prote
22	118	6.4	500	2 T32068	hypothetical prote
23	115.5	6.2	277	2 T25307	hypothetical prote
24	115.5	6.2	365	2 T25309	hypothetical prote
25	114	6.2	381	2 T15140	hypothetical prote
26	101	5.5	1123	2 AD2045	hypothetical prote
27	100	5.4	355	2 T34405	hypothetical prote
28	93	5.0	371	2 T32692	hypothetical prote
29	91	4.9	317	2 T33887	hypothetical prote

30 91 4.9 339 2 T34406 hypothetical prote  
31 91 4.9 401 2 B84828 hypothetical prote  
32 90 4.9 575 1 JC5432 glycoprotein 6-alp  
33 89.5 4.8 1462 1 S32437 pol polyprotein -  
34 89 4.8 451 2 E72398 hypothetical prote  
35 89 4.8 460 2 JC5137 beta-glucosidase (  
36 89 4.8 1489 2 S60416 DNA helicase YGL15  
37 87.5 4.7 420 2 D69252 L-carnitine dehydr  
38 87 4.7 1085 2 AH3171 DNA polymerase III  
39 86.5 4.7 483 1 A53595 allantoicase [EC 3  
40 86.5 4.7 805 2 E97717 virE4 protein prec  
41 85.5 4.6 252 2 AC2861 conserved hypothet  
42 85.5 4.6 276 2 B97638 hypothetical prote  
43 85 4.6 396 2 A84556 hypothetical prote  
44 85 4.6 718 2 G70978 probable copper-tr  
45 85 4.6 974 2 A72012 metalloproteinase,

ALIGNMENTS

RESULT 1  
A56098  
alpha(1,2)fucosyltransferase Sec2, long form - human  
C:Species: Homo sapiens (man)  
C>Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 21-Jul-2000  
R:Kelly, R.J.; Rouquier, S.; Giorgi, D.; Lennon, G.G.; Lowe, J.B.  
J. Biol. Chem. 270, 4640-4649, 1995  
A:Title: Sequence and expression of a candidate for the human Secretor blood group al  
tes with the non-Secretor phenotype.  
A:Reference number: A56098; MUID:95181460; PMID:7876235  
A:Accession: A56098  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-343 <KEL>  
A:Cross-references: GB:U17894; NID:g687618; PIDN:AAC24453.1; PID:g687619  
C:Genetics:  
A:Gene: GDB:FUT2; SE  
A:Cross-references: GDB:120619; OMIM:182100  
A:Map position: 19q13.3-19q13.3  
C:Keywords: glycoprotein; Golgi apparatus; transmembrane protein

Query Match	99.1%	Score 1832.5	DB 2	Length 343
Best Local Similarity	99.7%	Pred. No. 5.7e-149		
Matches 343	Conservative 0	Mismatches 0	Indels 1	Gaps 1
QY 1	MLVQMPFSPMAHFILEVFVTSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSOL 60			
Db 1	MLVQMPFSPMAHFILEVFVTSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSOL 60			
QY 61	RGWWTINATGRIGNOMGEYATLYALAKMNGRPAFTPAQMHSTLAPFRITLPLVLSATAS 120			
Db 61	RGWWTINATGRIGNOMGEYATLYALAKMNGRPAFTPAQMHSTLAPFRITLPLVLSATAS 120			
QY 121	RTPWQNYHLNDWMEERYRIIPGEYVRFTGYPCSWTFYHHLRQELIQEFTLHDHVREEAQ 180			
Db 121	RTPWQNYHLNDWMEERYRIIPGEYVRFTGYPCSWTFYHHLRQELIQEFTLHDHVREEAQ 179			
QY 181	KELRGLOVNGSRPGTFVGHVRRGDYVHVMPKWKGVADRRYLOQALDWFRRARYSSLIF 240			
Db 180	KELRGLOVNGSRPGTFVGHVRRGDYVHVMPKWKGVADRRYLOQALDWFRRARYSSLIF 239			
QY 241	VVTSNGMACRENIDTSHGDVVFAGDIGSGPAKDFALLTQCNHTTMTTGTGIAWAAYLT 300			
Db 240	VVTSNGMACRENIDTSHGDVVFAGDIGSGPAKDFALLTQCNHTTMTTGTGIAWAAYLT 299			
QY 301	GGDTIYLANTYLPDSPFLKIEKPEAFLEPWTGIAADLSPLLK 344			
Db 300	GGDTIYLANTYLPDSPFLKIEKPEAFLEPWTGIAADLSPLLK 343			

RESULT 2

B56392

beta-galactoside alphasal.2-fucosyltransferase II - rabbit

C:Species: Cryptotlagus cuniculus (domestic rabbit)

C:Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 05-Nov-1999

C:Accession: B56392

R:Hitoshi, S.; Kusunoki, S.; Kanazawa, I.; Tsuji, S.

J. Biol. Chem. 270, 8844-8850, 1995

A:Title: Molecular cloning and expression of two types of rabbit beta-galactoside alphasal.

A:Reference number: A56392; MUID:95238380; PMID:7721792

A:Accession: B56392

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-354 &lt;HIT&gt;

A:Cross-references: GB:X80225; NID:g854356; PIDN:CAA56512.1; PID:g854357

C:Keywords: transmembrane protein

Query Match 74.4%; Score 1375.5; DB 2; Length 354;  
Best Local Similarity 76.4%; Pred. No. 7e-110;  
Matches 253; Conservative 26; Mismatches 41; Indels 11; Gaps 3;

Qy 14 HFLFVFTVSTIFVQQLAKIQAMWELPVOIPVLASTSKALGPSQL--RGMTINAIGR 71

Db 29 YELTFIFVSVFVCHQLALVPAPWASRVVV-----PCHLPREGMTINAMGR 80

Qy 72 LGNMGEXATYALAKMNGRPAFIPAQMESTLAPIFRITLPLVLSATASRIPWQNYHLND 131

Db 81 LGNMGEXATYALAKENGREAYIPAQMESTLAPIFRITLPLVLSATASRIPWQNYHLND 140

Qy 132 WMEERYHIPGGEYVRFPGYSCSWTFYHHLRQELIQFTLHDHVRERAKFLRGLQVNGS 191

Db 141 WMEERYHIPV-PYVRLTGYPCSWTFYHHLRHEILREFTLHDHVRERAKFLRGLRNVGS 199

Qy 192 RPTFGVGVHVRGVDYVHMPKVGWADRYLQALDWRFRARYSSLIFVVTSGMAMCR 251

Db 200 RPTFGVGVHVRGVDYVHMPKVGWADRYLQALDWRFRAPAPPVFTVTSNGMAMCR 259

Qy 252 ENIDTSHGVDVVFAGDIEGSPAKDFALLTQCNHRTMTIGTFGIWAAYLTGDDTYLANYT 311

Db 260 ENIDASRGVDVVFAGDIEGSPAKDFALLTQCNHRTMTIGTFGIWAAYLTGDDTYLANYT 319

Qy 312 LPDSPFLKIFKPEAAFLPEWGIADLSPLL 342

Db 320 APDSFHLVFRPEAAFLPEWGIATNMGRAL 350

RESULT 3

A56392

beta-galactoside alphasal.2-fucosyltransferase I - rabbit

C:Species: Cryptotlagus cuniculus (domestic rabbit)

C:Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 05-Nov-1999

C:Accession: A56392

R:Hitoshi, S.; Kusunoki, S.; Kanazawa, I.; Tsuji, S.

J. Biol. Chem. 270, 8844-8850, 1995

A:Title: Molecular cloning and expression of two types of rabbit beta-galactoside alphasal.

A:Reference number: A56392; MUID:95238380; PMID:7721792

A:Accession: A56392

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-373 &lt;HIT&gt;

A:Cross-references: GB:X80226; NID:g854354; PIDN:CAA56513.1; PID:g854355

C:Keywords: transmembrane protein

Query Match 59.1%; Score 1093.5; DB 2; Length 373;  
Best Local Similarity 67.0%; Pred. No. 9.6e-86;  
Matches 203; Conservative 38; Mismatches 59; Indels 3; Gaps 2;

Qy 41 LPVQIPVLASTSKALGPSQLRGMNTINAIKRLGNMGEXATYALAKMNGRPAFIPAQM 100

Db 58 LPVTSFASNASSACGRPAAPSGIWTIHPDGRFGNMGQGYATLLAQLNGRRAPITLPA 117

Qy 101 STLAPIFRITLPLVLSATASRIPWQNYHLNDWMEERYHIPGGEYVRFPGYSCSWTF 160

Db 118 AALAPVFRITLPLVLAPEVNRRTSNKQLLDHWMSEEXSRL-EDPFLKFTGPGCSWTF 176

Qy 161 LRQELIQFTLHDHVRERAKFLRGLQV--NGSRPGTFVGVHVRGDDYVHMPKVGWGV 218  
Db 177 VREQIRREFTLHDHVRERAKFLRGLRGRGARPTFVGVHVRGDDYVHMPKVGWGV 236  
Qy 219 ADRLYLQOALDWRFRARYSSLIFVVTSGMAMCRNIDTSHGVDVVFAGDIEGSPAKDFAL 278  
Db 237 GDRAYLQOALDWRFRARHAPFVVTSGMAMCRNIDTSHGVDVVFAGDIEGSPAKDFAL 296  
Qy 279 LTQCNHRTMTIGTFGIWAAYLTGDDTYLANYTLPDSEFLKIFKPEAAFLPEWGIADL 338  
Db 297 LTQCNHRTMTIGTFGIWAAYLTGDDTYLANYTLPDSEFLKIFKPEAAFLPEWGIADL 356  
Qy 339 SPL 341  
Db 357 SPV 359

RESULT 4

A36047

galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) - human

C:Species: Homo sapiens (man)

C:Date: 16-Nov-1990 #sequence\_revision 16-Nov-1990 #text\_change 10-Sep-1997

C:Accession: A36047

R:Larsen, R.D.; Ernst, L.K.; Nair, R.P.; Lowe, J.B.

Proc. Natl. Acad. Sci. U.S.A. 87, 6674-6678, 1990

A:Title: Molecular cloning, sequence, and expression of a human GDP-L-fucose:beta-D-g

A:Reference number: A36047; MUID:90370848; PMID:2118655

A:Accession: A36047

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-365 &lt;LAR&gt;

A:Cross-references: GB:M35531; NID:g183887; PID:g306830

C:Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein

Query Match 58.4%; Score 1080.5; DB 2; Length 365;  
Best Local Similarity 59.7%; Pred. No. 1.2e-84;  
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;

Qy 12 MAHFLFVFTVSTIFVHVOOR-----LAKIQAMWELPVOIPVL-----ASPSK 53

Db 11 LAFLLVCVLVFIFFLIHHDQSPHGLGLSILCDPDRRLVTPPVVAFICLPGTAMGNASSC 70

Qy 54 ALGPSQLRGMNTINAIKRLGNMGEXATYALAKMNGRPAFIPAQMHTLAPIFRITL 113

Db 71 PQHPASLSSTVTVVNGRFGNMGQGYATLLAQLNGRRAPILPAMHAALAPVFRITL 130

Qy 114 LHSATASRIPWQNYHLNDWMEERYHIPGGEYVRFPGYSCSWTFYHHLRQELIQFTL 173

Db 131 LAPEYDSTPWRQLQHDWMSEEXADL-RDFTLKLSPGSCSWTFYHHLRQELIQFTL 189

Qy 174 HVREARAKFLRGLQV--NGSRPGTFVGVHVRGDDYVHMPKVGWADRYLQOALDWE 231

Db 190 HLREARAKFLRGLRGTGRPRFVGVHVRGDDYVHMPKVGWADRYLQOALDWE 249

Qy 232 RARYSSLIFVVTSGMAMCRNIDTSHGVDVVFAGDIEGSPAKDFALLTQCNHRTMTIGT 291

Db 250 RARHAPVFTVTVVNGRFGNMGQGYATLLAQLNGRRAPITLPA 309

Qy 292 FGIWAAYLTGDDTYLANYTLPDSEFLKIFKPEAAFLPEWGIADLSPL 341

Db 310 FGIWAAYLTGDDTYLANYTLPDSEFLKIFKPEAAFLPEWGIADLSPL 359

RESULT 5

S46494

galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) B - rat (fragment)

N:Alternate names: alpha-1,2-fucosyltransferase

C:Species: Rattus norvegicus (Norway rat)

C:Date: 26-Dec-1994 #sequence\_revision 23-Feb-1996 #text\_change 08-Oct-1999

A:Accession: S46494

R:Piau, J.P.; Labarriere, N.; Dabouls, G.; Denis, M.G.

Biochem. J. 300, 623-626, 1994

C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 28-Jul-2000  
C;Accession: T44328  
R;Yamasaki, S.; Shimizu, T.; Hoshino, K.; Ho, S.T.; Shimada, T.; Nair, G.B.; Takeda, Gene 237, 321-332, 1999  
A;Title: The genes responsible for O-antigen synthesis of *Vibrio cholerae* O139 are cl  
A;Reference number: 222749; MUID:9453293; PMID:10521656  
A;Accession: T44328  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-281 <YAM>  
A;Cross-references: EMBL:AB012957; NID:94115688; PIDN:BAA33632.1; PID:93721682  
A;Experimental source: strain O22  
C;Genetics:  
A;Note: wblA  
C;Superfamily: *Vibrio cholerae* hypothetical protein wblA

Query Match 10.5%; Score 195; DB 2; Length 281;  
Best Local Similarity 26.6%; Pred. No. 4.4e-09;  
Matches 77; Conservative 46; Mismatches 104; Indels 62; Gaps 14;

Qy 70 GRIGNOMGEYATLALAKNGRP-----AFTPAQMHSSTL-APIFRITLPVLH----- 115  
Db 8 GGLGNLFQYAVGRAIA:QYGVPLKLDYSAYKNYKLNHGYRLDQFNADIANEDEIFHL 67  
Qy 116 -----SATASRIPQNYHLNDWMEVEYRHI-----PPGEYVFTGYPGCSWTFYHHL 161  
Db 68 KGSSNRLSRILRLGW--LKKNTYYAEKORTIYDVSVFMQAPRYL--DGYWQNEQYFSOI 123  
Qy 162 ROEILQEFTHDHRVREAAQKFLRGLOVNGSRPGTVGVHVRGQDVHVMVKYKGVVADR 221  
Db 124 RAVLQELWPNQPLSINAQAQKIQOQTHA-----VSHVRGQDYLH-HPEIG---VLDI 174  
Qy 222 RYLOQALDFRARSYSLIFVTSNGMACKREN---IDTSHGQWVFGAGIEGSPAKDFAL 278  
Db 175 DYYKRAVDYIKKIEAPVFEVFNDAWCKQNFNIDSP---VFIED--TQTEIDDLML 228  
Qy 279 LQCNHTIMTIGTFGIAAYLTGG-DTYLANYT-----LPDS 315  
Db 229 MCQCGHNVANSFSSWAALNSVDKIVIAKPTWMAENPKYKWKVPDS 277

RESULT 8  
T20745  
hypothetical protein F11A5.5 - *Caenorhabditis elegans*  
C;Species: *Caenorhabditis elegans*  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000  
C;Accession: T20745  
R;Gardner, A.  
submitted to the EMBL Data Library, March 1997  
A;Reference number: Z19319  
A;Accession: T20745  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-363 <WIL>  
A;Cross-references: EMBL:292830; PIDN:CAB07352.1; GSPDB:GN00023; CESP:F11A5.5  
A;Experimental source: clone F11A5  
C;Genetics:  
A;Gene: CESP:F11A5.5  
A;Map position: 5  
A;Introns: 42/1; 55/2; 136/3; 175/3; 210/2; 243/1; 299/1; 330/3  
C;Superfamily: *Caenorhabditis elegans* hypothetical protein C14C6.3

Query Match 9.8%; Score 181.5; DB 2; Length 363;  
Best Local Similarity 23.3%; Pred. No. 8.6e-08;  
Matches 87; Conservative 54; Mismatches 116; Indels 117; Gaps 20;

Qy 16 ILFVFTVSTIFHVQORLAKIQAMWEL-PVQIPVLASTKALGPSQLRGWMTNAI----- 59  
Db 33 LEWMILVYSIF-----LIRAMHEEDPMEVP-----YPPAQINFTSRKYSNYA 77  
Qy 70 --GRIGNOMGEYATLALAKMNGR-PAFIPAQM-----STIAPIFRITLPVLHSAT 118  
Db 78 SSSRLGNHLFELASVLSISRELQRFVTFIENCYHEKMKWEDSNTLIPGLMNHFLINGSV 137

A;Title: Evidence for two distinct alpha-(1,2)-fucosyltransferase genes differentially e  
A;Reference number: S46493; MUID:94280382; PMID:8010942  
A;Accession: S46494  
A;Molecule type: mRNA  
A;Residues: 1-159 <PIA>  
A;Cross-references: EMBL:L26010; NID:9414816; PIDN:AAB41514.1; PID:9414817  
C;Genetics:  
A;Gene: FTB  
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 33.7%; Score 623; DB 2; Length 159;  
Best Local Similarity 88.5%; Pred. No. 5.3e-46;  
Matches 116; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 213 VKKGWVADRRYLQALDFRARSYSLIFVTSNGMACKRENIDTSHGQDVVFGAGIEGSP 272  
Db 2 VKKGWVADRGYLEKALDFRARSYSPVFTVTSNGMACKRENINASRGDGVFAGNGIEGSP 61  
Qy 273 AKDFALLTQCNTHTMTIGTFGIWAAYLTGSPITLYANLYTLPDSPFLKTFKPEAAFLPWT 332  
Db 62 AKDFALLTQCNTHTMTIGTFGIWAAYLAGSDTIYLYANLYTLPDSPFLKTFKPEAAFLPEW 121  
Qy 333 GIAADLSPLK 343  
Db 122 GIPADLSPLK 132

RESULT 6  
S51582  
galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) - rat  
C;Species: *Rattus norvegicus* (Norway rat)  
C;Date: 07-May-1995 #sequence\_revision 17-Nov-1995 #text\_change 08-Oct-1999  
C;Accession: S51582; S46493  
R;Plau, J.P.; Labarriere, N.; Dabouis, G.; Denis, M.G.  
Submitted to the EMBL Data Library, November 1993  
A;Reference number: S51582  
A;Accession: S51582  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-142 <PIA>  
A;Cross-references: EMBL:L26009; NID:9414814; PIDN:AAB41514.1; PID:9554438  
R;Plau, J.P.; Labarriere, N.; Dabouis, G.; Denis, M.G.  
Biochem. J. 300, 623-626, 1994  
A;Title: Evidence for two distinct alpha-(1,2)-fucosyltransferase genes differentially e  
A;Reference number: S46493; MUID:94280382; PMID:8010942  
A;Accession: S46493  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 9-135 <PI2>  
A;Cross-references: EMBL:L26009  
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 32.1%; Score 593; DB 2; Length 142;  
Best Local Similarity 74.6%; Pred. No. 1.7e-43;  
Matches 106; Conservative 16; Mismatches 18; Indels 2; Gaps 1;

Qy 152 PCSWTFYHHLRQFIHQEFTLHDHVRVREAAQKFLRGLOVNG--GSRPGTFVGVHVRGQDVHV 209  
Db 1 PCSWTFYHHLRQIRREFTLHDHVRVREAAQKFLRGLOVNG--GSRPGTFVGVHVRGQDVHV 60  
Qy 210 MPKWKGVVADRRYLQALDFRARSYSLIFVTSNGMACKRENIDTSHGQDVVFGAGIEG 269  
Db 61 MPNRKGVVGDRAYLQKAMDWRARHKDPFVFTVTSNGMACKRENIDTSHGQDVVFGAGIEG 120  
Qy 270 GSPAKDFALLTQCNTHTMTIGT 291  
Db 121 GTPGKDFALLTQCNTHTMTIGT 142

RESULT 7  
T44328  
hypothetical protein wblA [imported] - *Vibrio cholerae*  
C;Species: *Vibrio cholerae*

QY 119 ASRIPQNVH-----LNDWMEERYSHIPPGYVVRFTGYPGWSYFHHLRQBILOE 168  
 Db 136 PSSVRKVKHQKCCFFDSSLNDNYDEYHLHTGTHTYO-----SKYFESHMRNELIGY 190  
 QY 169 FTLHDVREAAQKFLRGLGVNSRGTFTYG-VHVRGDYVHVMPKVGK-VADRRYLAQ 226  
 Db 191 LKTTENTYMDLEK-----SGENTFTTCVHVRGDFLRV-----CFHVADENFIRS 235  
 QY 227 ALDWFPRYSSLIFFVTSNGMAWCHRENDTSHGDVVVFAGDGE----- 269  
 Db 236 SUNLISROVAK-----RANTAT-----VFFGDDYEFMDSLRNRSTKINAFVS 277  
 QY 270 -GSPAKDFALL-TOCNHITMTI--GTGFWAAAYLGGDTIYLIANTYLPDSPFLK----- 319  
 Db 278 QNSPADLLYAKSNCDDVLLTAHSTFTGWMGYTSKGRNVY---YT--DIQFTKDWILET 332  
 QY 320 -IFKPEAAFLPEWT 332  
 Db 333 GEFISEDYLPWHT 346

RESULT 9  
 H87911  
 protein B0205.4 [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
 C:Accession: H87911  
 R:anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
 A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.ele  
 A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A/Accession: H87911  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-443 <STO>  
 A/Cross-references: GB:Chr.I; PIDN:AAC16988.1; PID:93150470; GSPDB:GN00019; CESP:B0205.4  
 A/Note: contains weak similarity to fucosyltransferases  
 C:Genetics:  
 A:Gene: B0205.4  
 A/Map position: 1

Query Match 9.5%; Score 176; DB 2; Length 443;  
 Best Local Similarity 20.4%; Pred. No. 3.3e-07;  
 Matches 89; Conservative 63; Mismatches 145; Indels 140; Gaps 17;

QY 15 FILFVFTVS-----TIFHVQORLAKI-QAMWELPV-----QIPVLA 49  
 Db 29 YILFLENLSKTRLTPEERATYMKKNVYQSGNEATLDCPW--PISNVENVINRVSYLE 86  
 QY 50 STSKA-----LGPSQLRGWMTINAGRLGNMGCEYATLYALAKMNGRPAPFIPAOMHSTLA 104  
 Db 87 NSCKRIFSDFGYSQ-----GIGNLMPQVAGLLSIARETSILLIPST--TLTLR 133  
 QY 105 PTF-----RTLPVLHSATASRPWON-----VHLNDWMEERYSHIPPGY 145  
 Db 134 RAFDFETTFNDSTQFVGEDLSQLAEDLNASKTLITLSCAYRNLSILFNDKRIIE----- 189  
 QY 146 VFTGYPGWSYFHHLRQBILOE-FTLHDVRE----- 177  
 Db 190 -RIDGYFQNFYFHPDSOKIVKLFTEFMDPVKRVKRSFVNIYWNHTPHNRKKPEKSTVS 248  
 QY 178 -----EAKQFLR-----GLQVNGSR-----PGTFVGVHVR 202  
 Db 249 IFSVTFQURVDFLENVGISTVRNARIETNVANDQALELPEEDAFAKTMVGVHLR 308  
 QY 203 RGDYVHV--MPKWKGVADRYLQALDWPFRARYSSL:FVYTSNGMAWCRENIDTSHGD 260  
 Db 309 HGMDSMSNRNIHGHVDTPYIYKRAQQISKIYENAFITCSDNVAVARENKLKRET 368  
 QY 261 VYFAGDIEGSPAKDFALLTQCNHITMTITGTFGIWAAYL---TGDDTIYLANITLPDSPFF 317

Db 369 LHFF---CGCPREVDMAILKSCDSVLIISTGTFGWSAYLNVSAPDVIYKHPAPGSMV 425  
 QY 318 LKIFKPEAAFLPEWTGI 334  
 Db 426 EXMINKTEYFLAKSWTAL 442

RESULT 10  
 T32309  
 hypothetical protein F31F4.11 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 20-Jun-2000  
 C:Accession: T32309  
 R:Blanchard, M.; Kramer, J.; Elliott, G.; Twyman, B.  
 submitted to the EMBL Data Library, September 1997  
 A/Description: The sequence of C. elegans cosmid F31F4.  
 A/Reference number: Z21149  
 A/Accession: T32309  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-395 <BLA>  
 A/Cross-references: EMBL:AF024503; PIDN:AAB70387.1; GSPDB:GN00023; CESP:F31F4.11  
 A/Experimental source: strain Bristol N2; clone F31F4  
 C:Genetics:  
 A:Gene: CESP:F31F4.11  
 A/Map position: 5  
 A/Introns: 22/3; 98/2; 179/3; 214/2; 248/1; 283/3; 329/1; 360/3  
 C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 8.4%; Score 155; DB 2; Length 395;  
 Best Local Similarity 20.4%; Pred. No. 1.8e-05;  
 Matches 77; Conservative 47; Mismatches 110; Indels 144; Gaps 18;

QY 18 FVFTVSTIFHVQORLAKIQAMWELPVQIP---VLASTSKALGSPQLRGMTWINAIGRLGN 74  
 Db 80 YFSAETLF-----SKNSNFWLSGIPYPALINFTSSQVLSNLASPC-----QLGN 127  
 QY 75 QMGVATLYALAK-MNGRPA-FIPAQMHSILAPFRTLPVLHSATASRPWQNYHLNDW 132  
 Db 128 HIFELASLYGLSKHLNRTPAFFESGYHKNMDSIRSTMPGL----- 169  
 QY 133 MEERYSHIPPGYVVRFTGYPGWSYFHHLRQBILOEFTLHDVREAAQKFLRGLQVNGSR 192  
 Db 170 -----IGKYAIPDG--SSMKYFPNMRRELI-----SFLNTSVQNF--GI-LPISN 209  
 QY 193 PGTFVG-VHVRGDYVHVMPKVGKGVYVADRYLQALDWPFRARYSSLIFVYTSNGMAWCR 251  
 Db 210 NNTHVSCVHSRRGDFEVM-----NFYATDPKMKNAVKELKE-----K 247  
 QY 252 ENIDTSHGDVVVFAGDGI-----EGSPAKDFALLT-QCNHT 285  
 Db 248 ENYSKNQKIVLFGDDFKPMRNLFSEAKYSTDAHESVEYIYSONSAIDDFLYSAYNCDEV 307  
 QY 286 IMP-----IGTFGIWAAYLTGGDTIYLA-----NYLTPD 314  
 Db 308 LITGNAHQPTSMHVANFNLDITPRSTFGWNLGYFGSKGVYVLDIKHARDVYDNGEIAE 367  
 QY 315 SPLKIFKPEAAFLPEWT 332  
 Db 368 NDF-----FLPHWT 376

RESULT 11  
 T31916  
 hypothetical protein C17A2.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jul-2000  
 C:Accession: T31916  
 R:Saumons, L.; Wohldmann, P.  
 submitted to the EMBL Data Library, July 1997  
 A/Description: The sequence of C. elegans cosmid C17A2.  
 A/Reference number: Z21098

125334 hypothetical protein T26H5.8 - *Caenorhabditis elegans*

C:Species: Caenorhabditis elegans  
C:date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jul-2000  
C:Accession: T25334; T26600  
R:Gardner, A.  
submitted to the EMBL Data Library, November 1996

A:Reference number: Z20017  
A:Accession: T25334  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-392 <WIL>  
A:Cross-references: EMBL:T282056; PIDN:CA04857.1; GSPDB:GN00023; CESP:T26H5.8  
A:Experimental source: clone T26H5  
R:Wall, M.

submitted to the EMBL Data Library, October 1998  
A:Reference number: Z20243  
A:Accession: T26600  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-392 <WII2>  
A:Cross-references: EMBL:AL032620; PIDN:CAA21488.1; CESP:T26H5.8  
A:Experimental source: clone Y3E3A  
C:Genetics:  
A:Gene: CESP:T26H5.8  
A:Map position: 3  
A:Intronous: 7/1; 32/3; 55/1; 109/3; 148/3; 187/2; 222/L; 263/L; 333/L; 363/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein CL4C6.3

Query Match            7.4%   Score 137.5; DB 2; Length 392;  
Best Local Similarity 21.0%; Pred.No. 0.00055;  
Matches 72; Conservative 56; Mismatches 114; Indels 101; Gaps 18;

QY 72 LGNQMGQEYATIALAKMGRPAFIPA-----QMHSITAPIF-----RTLPVLHSDATA 119  
| | | | :  
Db 54 LGLRFELISLGIASTLORRAVINATPNVNSHLHQTMQLPKLVQEFELRVIPESVV 113  
| | | | :  
QY 120 SR--IPWNQNYHLND----WMEEYRH-IPPGEYVRFTPGCSWTFFHYHLROELLQEFFTLH 172  
:  
Db 114 ARQQINWSRCICFDDEPAKVLDISNGHVLDGHYPQ-----SFKFHIRIKRWMAKP 167  
:  
QY 173 DHVREAAOKFLRGCVNGSRGTGVGVHVRGDY-----VHYMPKVWGVAADRRYLQAAL 228  
:  
Db 168 ELASKAEKL I-----PFSVKDNFIICPIHRRGDFTDLGH-QP-----SDPTETRAAT 215  
:  
QY 229 DWFARYS-----LI FVVTSGMA-----WCRENIDTSH-GDVVFAGDG- 267  
:  
Db 216 DFLEVKYQKHRRHTVVVVGNDMSFAHTVFEGEKSFFFCFCKKYEILLHRVGNISFALNP 275  
:  
QY 268 -----IEGSPA-----KDFALL-----TCNNHNTI---I 289  
| | | | :  
Db 276 ATALNYLTPEYSPTYDVRSNRFLAATLKHFVOILTPOSTPEIDLAFSRFCDAITLTGPS 335  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
QY 290 GTFGIWAYLTGGDT-TYLANLYTPDSFPFLKFKEPA AFLPW 331  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 336 STFGWLWLSLA KSAKVYYRDILETKDGVINDMKVEDFYYPEW 378

RESULT 15  
T02798

hypothetical protein L549.10 [imported] - Leishmania major (strain Friedlin)  
C:Species: Leishmania major  
C:date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 19-May-2000  
C:Accession: B81456; T02798  
R:Myler, P.J.; Audleman, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.;  
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999  
A:title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-co  
A:Reference number: B81455; MUID:99178987; PMID:10077609  
A:Accession: B81456  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-348 <PYL>  
A:Cross-references: GB:AE001274; NID:g32464850; PIDN:AAC24622.1; PID:g2978459; GSPDB:GN00  
A:Experimental source: strain MHOW/IL/81/Friedlin  
C:Genetics:

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:53:11 ; Search time 10.0683 Seconds

(without alignments)  
1417.108 Million cell updates/sec

Title: US-10-040-863-11

Perfect score: 1850

Sequence: 1 MLVYMPFSPFMAHFLVF.....AAFLPEWTGTAADLSPLIKH 344

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	1832.5	99.1	343	FUT2_HUMAN	Q10981 h galactosi
2	1528	82.6	340	FUT2_PIG	Q10982 sus scrofa
3	1510	81.6	344	FUT2_BOVIN	Q28113 bos taurus
4	1375.5	74.4	354	FUT2_RABIT	Q10983 oryctolagus
5	1299.5	70.2	368	FUT2_MOUSE	P97353 mus musculus
6	1100.5	59.5	376	FUT1_RAT	Q10980 rattus norv
7	1093.5	59.1	373	FUT1_RABIT	Q10979 oryctolagus
8	1083.5	58.6	376	FUT1_MOUSE	Q09160 mus musculus
9	1080.5	58.4	365	FUT1_HUMAN	P19326 homo sapien
10	1079.5	58.4	365	FUT1_PIG	Q29043 sus scrofa
11	623	33.7	159	FUT2_RAT	Q10984 rattus norv
12	155	8.4	365	YK07_CAEEL	P34302 caenorhabdi
13	92	5.0	575	FUT8_PIG	P79282 s alpha-(1,
14	90	4.9	575	FUT8_HUMAN	Q9bys5 h alpha-(1,
15	90	4.9	575	FUT8_MOUSE	Q9bys2 m alpha-(1,
16	89	4.8	1489	YGP0_YEAST	P53115 saccharomyc
17	86.5	4.7	484	ALN_RANCA	P40757 rana catesb
18	85	4.6	323	MIAA_RALSO	Q8xb0 ralstonia s
19	85	4.6	718	CTPC_MYCTU	P96875 mycobacteri
20	84.5	4.6	630	S6A4_CAVPO	C35899 cavia porce
21	84	4.5	433	CBPA_AMOGA	Q02350 anopheles g
22	83.5	4.5	660	FBUB_ECOLI	P06972 escherichia
23	83	4.5	328	GLK_NEIMA	Q9jx53 neisseria m
24	82.5	4.5	1048	SBCC_ECOLI	P13458 escherichia
25	82	4.4	575	FUT8_BOVIN	Q9n0w2 b alpha-(1,
26	82	4.4	1211	ATC6_SCHPO	Q14072 schizosacch
27	82	4.4	1928	LPH_RAT	Q02401 rattus norv
28	81.5	4.4	476	NOEA_RHIME	Q52892 rhizobium m
29	81.5	4.4	537	FUT6_ARATH	Q9xi80 arabidopsis
30	81.5	4.4	630	S6A4_MOUSE	C60857 mus musculus
31	81.5	4.4	630	S6A4_RAT	P31652 rattus norv
32	81.5	4.4	1279	SCAP_HUMAN	Q12770 homo sapien
33	80.5	4.4	874	BCSA_SALTY	Q93in2 salmoneila

RESULT 1	FUT2_HUMAN	STANDARD;	PRT;	343 AA.
34	80	4.3	990	1 KGPL_PICPA
35	79.5	4.3	353	1 FES_YEREN
36	79	4.3	454	1 NOE2_HUMAN
37	79	4.3	776	1 PSAB_HETTR
38	79	4.3	1461	1 TOP2_CANAL
39	78.5	4.2	645	1 LIPL_PHOIU
40	78	4.2	470	1 PABF_BACSU
41	78	4.2	890	1 POL2_BAYMG
42	78	4.2	963	1 TRES_THETH
43	78	4.2	2294	1 YGF2_ARATH
44	78	4.2	2479	1 POLN_RRVN
45	77.5	4.2	395	1 HMPA_ERWCH
ALIGNMENTS				
AC	Q10981;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)Ft 2)			
DE	(Fucosyltransferase 2) (Secretor blood group alpha-2-fucosyltransferase) (Secretor factor) (Se) (SE2).			
GN	FUT2 OR SEC2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95181460; PubMed=7876235;			
RA	Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;			
RT	"Sequence and expression of a candidate for the human Secretor blood group alpha(1,2)fucosyltransferase gene (FUT2). Homozygosity for an enzyme-inactivating nonsense mutation commonly correlates with the non-secretor phenotype."			
RL	J. Biol. Chem. 270:4640-4649(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND VARIANT PHE-140.			
RX	MEDLINE=96199252; PubMed=8621666;			
RA	Kudo T., Iwasaki H., Nishihara S., Shinya N., Ando T., Narimatsu I., Narimatsu H.;			
RT	"Molecular genetic analysis of the human Lewis histo-blood group system. II. Secretor gene inactivation by a novel single missense mutation A385T in Japanese nonsecretor individuals."			
RL	J. Biol. Chem. 271:9830-9837(1996).			
RN	[3]			
RP	SEQUENCE FROM N.A., AND VARIANT PHE-140.			
RX	MEDLINE=97363210; PubMed=9219535;			
RA	Koda Y., Soejima M., Wang B., Kimura H.;			
RT	"Structure and expression of the gene encoding secretor-type galectoside 2-alpha-L-fucosyltransferase (FUT2)."			
RL	Eur. J. Biochem. 246:750-755(1997).			
RN	[4]			
RP	VARIANTS VAL-25; CYS-138 AND ASN-172.			
RX	MEDLINE=98431007; PubMed=9760207;			
RA	Liu Y., Koda Y., Soejima M., Pang H., Schlaphoff T., du Toit E.D., Kimura H.;			
RT	"Extensive polymorphism of the FUT2 gene in an African (Xhosa) population of South Africa."			
RL	Hum. Genet. 103:204-210(1998).			
CC	!- FUNCTION: CREATES A SOLUBLE PRECURSOR OLIGOSACCHARIDE FUC-ALPHA-((1,2)GALBETA-) CALLED THE H ANTIGEN WHICH IS AN ESSENTIAL SUBSTRATE FOR THE FINAL STEP IN THE SOLUBLE A AND B ANTIGEN SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE THE SAME ACCEPTOR SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.			
CC	!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP + alpha-L-fucosyl-1,2-beta-D-galactosyl-R.			

Q92448 pictia past  
Q58555 yersinia en  
Q95897 homo sapien  
Q9xgv2 heterocapsa  
P87078 candida alb  
P40601 photorhabdu  
P28820 bacillus su  
Q01365 bailey yell  
Q06458 thermus the  
P56786 arabidopsis  
P13887 ross river  
Q47266 erwiania chr





```

FT DOMAIN 29 340 (POTENTIAL).
FT CARBOHYD 185 195 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 279 279 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 158 158 R -> H (IN REF. 3).
FT SEQUENCE 340 AA; 38987 MW; 0629F1C04FC206AD CRC64;

Query Match 82.6%; Score 1528; DB 1; Length 340;
Best Local Similarity 82.0%; Pred. No. 1.le-123;
Matches 282; Conservative 24; Mismatches 34; Indels 4; Gaps 2;

QY 1 MLVQMFSPMAHIFLVFTVSTIFHVOQLAKIQAMWELPVOIPVLASTSKALGPSQL 60
DB 1 MLSQASFFPTGPFILFVFTASTIFHLQRMVKIQITWELQM--VTQVTESPPSQL 57
QY 61 RGMWNTINAIGRLGNOMGEYATLYALAKMNGRPAFIPQAMHSTLAPIFRITLPVLHSATAS 120
DB 58 KGMWNTINAIGRLGNOMGEYATLYALAKMNGRPAFIPQAMHSTLAPIFRITLPVLHASTAR 117
QY 121 RIPQWNTHLNDWMEEEYRHPGGEYVFTGPGCSWTYFHHLRQELQEFTHDHRVREAA 180
DB 118 RIPQWNTHLNDWMEEEYRHPGGEYVFTGPGCSWTYFHHLRQELQEFTHDHRVREAA 176
QY 181 KFLRLGVNGSRPGTFVGVHVRGDYVHVPKVKGVVADRYLQOALDWFPRARYSLIF 240
DB 177 DFLRLGVNGSRPGTFVGVHVRGDYVHVPKVKGVVADRYLQOALDWFPRARYSPVF 236
QY 241 VVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTCNHTIMTIGFTGWAAYLT 300
DB 237 VVSSNGMAWCRENINASRGDDVVFAGDIEGSPAKDFALLTCNHTIMTIGFTGWAAYLA 296
QY 301 GGDRIYLYANTLPDPSFLKIFKPEAAFLPEWTGTAADLSPLKH 344
DB 297 GGETIYLYANTLPDPSFLKIFKPEAAFLPEWTGTAADLSPLKH 340

RESULT 3
FUT2_BOVIN STANDARD; PRT; 344 AA.
AC Q28113;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (GDP-L-fucose:beta-D-
DE D-galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)Ft 2)
DE (Fucosyltransferase 2).
GN FUT2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Petit J.M.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X99620; CAA67931.1; -

```

```

DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Golgi stack; Signal-anchor.
FT DOMAIN 1 7 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 8 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 29 344 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 344 AA; 39320 MW; F35AC33F1B7B9F68 CRC64;

Query Match 81.6%; Score 1510; DB 1; Length 344;
Best Local Similarity 82.3%; Pred. No. 4.le-122;
Matches 284; Conservative 14; Mismatches 45; Indels 2; Gaps 2;

QY 1 MLVQMFSPMAHIFLVFTVSTIFHVOQLAKIQAMWELPVOIPVLASTSKALGP-SQ 59
DB 1 MFSQTQTEFFPTAPFILFVFTASTIFHLQRLKMQPTWELEAPATMETPSRPPQEPQ 60
QY 60 LRGWNTINAIGRLGNOMGEYATLYALAKMNGRPAFIPQAMHSTLAPIFRITLPVLHSATA 119
DB 61 LRGWNTINAIGRLGNOMGEYATLYALAKMNGRPAFIPQAMHSTLAPIFRITLPVLHDATA 120
QY 120 SRIPQWNTHLNDWMEEEYRHPGGEYVFTGPGCSWTYFHHLRQELQEFTHDHRVREAA 179
DB 121 RSVQWNTHLNDWMEEEYRHPGGEYVFTGPGCSWTYFHHLRQELQEFTHDHRVREAA 179
QY 180 QKFLRLGVNGSRPGTFVGVHVRGDYVHVPKVKGVVADRYLQOALDWFPRARYSLI 239
DB 180 QNFLRLGVNGSRPGTFVGVHVRGDYVHVPKVKGVVADRYLQOALDWFPRARYSAPI 239
QY 240 FVVTSGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTCNHTIMTIGFTGWAAYLT 299
DB 240 FVSSNGMAWCRENINASRGDDVVFAGDIEGSPAKDFALLTCNHTIMTIGFTGWAAYLT 299
QY 300 TGGDRIYLYANTLPDPSFLKIFKPEAAFLPEWTGTAADLSPLKH 344
DB 300 AGGETIYLYANTLPDPSFLKIFKPEAAFLPEWTGTAADLSPLKH 344

RESULT 4
FUT2_RABIT STANDARD; PRT; 354 AA.
AC Q10383;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (Secretor blood
DE group alpha-2-fucosyltransferase) (GDP-L-fucose:beta-D-
DE galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)Ft 2)
DE (Fucosyltransferase 2).
GN FUT2 OR SECL OR RFT-II.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Hitoshi S., Kusunoki S., Kanazawa I., Tsuji S.;
RT "Molecular cloning and expression of two types of rabbit beta-
RT galactoside alpha 1,2-fucosyltransferase."
RL J. Biol. Chem. 270:8844-8850(1995).
CC -!- FUNCTION: CREATES A SOLUBLE PRECURSOR OLIGOSACCHARIDE FUC-ALPHA
CC ((1,2)GALBETA-) CALLED THE H ANTIGEN WHICH IS AN ESSENTIAL
CC SUBSTRATE FOR THE FINAL STEP IN THE SOLUBLE A' AND B ANTIGEN
CC SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE THE SAME ACCEPTOR
CC SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +

```

```
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -1- PATHWAY: Glycosylation.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -1- TISSUE SPECIFICITY: SALIVARY AND LACTATING MAMMARY GLANDS.
CC -1- MISCELLANEOUS: THERE ARE THREE GENES (FUT1, FUT2 AND FUT3) WHICH
CC ENCODE GALACTOSIDE 2-L-FUCOSYLTRANSFERASE IN RABBIT. THEY ARE
CC EXPRESSED IN A TISSUE-SPECIFIC MANNER.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X80225; CAA56512.1; -.
CC InterPro: IPR002516; GT_11.
CC Pfam: PF01531; Glyco_transf_11; 1.
CC Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
CC Signal-anchor; Golgi stack.
CC DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 23 43 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC (POTENTIAL).
CC DOMAIN 44 354 LUMENAL, CATALYTIC (POTENTIAL).
CC CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 317 317 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 354 AA; 40035 MW; 1E2B831F9DA6C0B4 CRC64;
CC -----
Query Match 74.4%; Score 1375.5; DB 1; Length 354;
Best Local Similarity 76.4%; Pred. No. 1.4e-110;
Matches 253; Conservative 26; Mismatches 41; Indels 11; Gaps 3;
Qy 14 HFILFVTVSTFHFVQORLAKIQAMWELPQVIPVLASTSKALGPSQ--RCMTWINAIGR 71
Db 29 YLFETTFVSTFHFVQORLAKIQAMWELPQVIPVLASTSKALGPSQ--RCMTWINAIGR 80
Qy 72 LGNOMGEYATLYALAKMNGRPAFIPAQMHSHTLAFITRTPLVLSHATSRIPWNYHLND 131
Db 81 LGNOMGEYATLYALAKMNGRPAFIPAQMHSHTLAFITRTPLVLSHATSRIPWNYHLND 140
Qy 132 WNEEYRHPGEEVYRFTGYPCSWTFYHLLRQEILOEFTLHDHVRERAAQAFRLGLRVNGS 191
Db 141 WNEEYRHPGEEVYRFTGYPCSWTFYHLLRQEILOEFTLHDHVRERAAQAFRLGLRVNGS 199
Qy 192 PGTFTGVGVHVRGDDYVHVMKPKVGVVADRYVLOQALDWERARYSSLIFFVTSNGMWR 251
Db 200 PGTFTGVGVHVRGDDYVHVMKPKVGVVADRYVLOQALDWERARYSSLIFFVTSNGMWR 259
Qy 252 ENIDTSHGDVVYFAGDIEGSPAKDFALLTQCNRHMTITGTCINAAAYLTGGDTIYLANYT 311
Db 260 ENIDTSHGDVVYFAGDIEGSPAKDFALLTQCNRHMTITGTCINAAAYLTGGDTIYLANYT 319
Qy 312 LPDGFPLKIFKPERAFPEWGTGIADLSPL 342
Db 320 APDGFPLKIFKPERAFPEWGTGIADLSPL 350
RESULT 5
FUT2_MOUSE STANDARD; PRT; 368 AA.
AC P97353;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (GDP-L-fucose:beta-
DE D-galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)FUT 2)
DE (Fucosyltransferase 2) (FUT-III).
GN FUT2 OR SECL.
OS Mus musculus (Mouse).
```

```
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC NCBI_TaxID=10090;
CC [1]
CC RN SEQUENCE FROM N.A.
CC STRAIN=ICR;
CC Hitooshi S., Kusunoki S., Kanazawa I., Tsuji S.;
CC "Molecular cloning and expression of a mouse GDP-L-Fucose: beta-D-
CC galactoside 2-alpha-L-Fucosyltransferase.";
CC Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC [2]
CC RN SEQUENCE FROM N.A.
CC STRAIN=ICR; TISSUE=Gastrointestinal tract;
CC Lin B., Hayashi Y., Saito M., Sakakihara Y., Yanagisawa M.,
CC Iwamori M.;
CC "Molecular cloning and expression of a GDP-L-fucose:beta-D-galactoside
CC 2-alpha-L-fucosyltransferase in murine gastrointestinal tract.";
CC Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -1- PATHWAY: Glycosylation.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).
CC -1- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND SECL) WHICH ENCODE
CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y09882; CAA71008.1; -.
CC EMBL: AF113532; AAD25351.1; -.
CC MGD: MGI109374; Fut2.
CC InterPro: IPR002516; GT_11.
CC Pfam: PF01531; Glyco_transf_11; 1.
CC Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
CC Signal-anchor; Golgi stack.
CC DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 21 41 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC (POTENTIAL).
CC DOMAIN 42 368 LUMENAL, CATALYTIC (POTENTIAL).
CC CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 315 315 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 368 AA; 41464 MW; 4093E853EB37303B CRC64;
CC -----
Query Match 70.2%; Score 1299.5; DB 1; Length 368;
Best Local Similarity 73.6%; Pred. No. 4.8e-104;
Matches 245; Conservative 30; Mismatches 43; Indels 15; Gaps 4;
Qy 11 PNAHFLP--VFTSTFHFVQORLAKIQAMWELPQVIPVLASTSKALGPSQ---LRGMWT 65
Db 22 PLSTFLEFVIFVSTFHFVQORLAKIQAMWELPQVIPVLASTSKALGPSQ---LRGMWT 72
Qy 66 INATGRLGNQMGVATLYALAKMNGRPAFIPAQMHSHTLAFITRTPLVLSHATSRIPWQ 125
Db 73 IRVKGRLGNQMGVATLYALAKMNGRPAFIPAQMHSHTLAFITRTPLVLSHATSRIPWQ 132
Qy 126 NYHLNDWMEERYRHIPGEEVYRFTGYPCSWTFYHLLRQEILOEFTLHDHVRERAAQAFRLG 185
Db 133 NYHLNDWMEERYRHIPGEEVYRFTGYPCSWTFYHLLRQEILOEFTLHDHVRERAAQAFRLG 191
Qy 196 LOVNGSRGCTFVGIVHVRGDDYVHVMKPKVGVVADRYVLOQALDWERARYSSLIFFVTSN 245
Db 192 LOVNGSRGCTFVGIVHVRGDDYVHVMKPKVGVVADRYVLOQALDWERARYSSLIFFVTSN 251
Qy 246 GMAWGRNIDTSHGDVTFAGDIEGSPAKDFALLTQCNRHMTITGTCINAAAYLTGGDTI 305
Db 246 GMAWGRNIDTSHGDVTFAGDIEGSPAKDFALLTQCNRHMTITGTCINAAAYLTGGDTI 305
```

```
Db 252 DMWCRKSITASRGDVFAGNGLOGSPAKDIALMQCNHTVITLCTFGIWAAYLTGGDTV 311
QY 306 YLANETLPDSEFLKIFKPEAAFLPEWTGIAADL 338
Db 312 YLANETQNSPFHTVFKPEAAFLPEWVGIAADL 344

RESULT 6
FUT1_RAT
ID FUT1_RAT STANDARD; PRT; 376 AA.
AC Q10980;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-
D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)
DE (Fucosyltransferase 1).
GN FUT1 OR FTA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon cancer;
RA Soejima M., Wang B., Koda Y., Kimura H.;
RT "Two distinct rat GDP-L-fucose:b-D-galactoside 2-alpha-L-
fucosyltransferase genes";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 169-310 FROM N.A.
RC STRAIN=BDIX;
RX MEDLINE=94280382; PubMed=8010942;
RA Piau J.-P., Labarriere N., Dabouis G., Denis M.G.;
RT "Evidence for two distinct alpha(1,2)-fucosyltransferase genes
differentially expressed throughout the rat colon.";
RL Biochem. J. 300:623-626(1994).
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
FORM IN TRANS CISTERNAE OF GOLGI.
CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF
MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB015637; BAA31130.1; -
CC EMBL; AB006137; BAA21741.1; -
CC EMBL; L26009; AAB41514.1; -
CC InterPro: IPR002516; GT_11.
CC Pfam: PF01531; Glyco.transf.11; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
FT SIGNAL-anchor; Golgi stack.
FT DOMAIN 1 12
FT TRANSMEM 13 29
FT CYTOPLASMIC (POTENTIAL).
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT FT 306
FT LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 64
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 302
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 376 AA; 42416 MW; C350C737C758B7F8 CRC64;

Query Match 59.5%; Score 1100.5; DB 1; Length 376;

Best Local Similarity 67.0%; Pred. No. 5.7e-87;
Matches 207; Conservative 36; Mismatches 59; Indels 7; Gaps 4;
QY 37 AMWELPVOIPVLASTSKALGPSQ--LRGMWTNIALGRIGNQMGCEYATLYALAKMNGRPAF 94
||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 55 AMVCLP--VPSNASSGSPSCPEQLLSGTTWITPGRTFGNQMGQYATLLALAAQLNGRAAF 112
||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 95 IPAQMSHILAPIFRITLPLVILHSATASRIPQWNYHLNDWMEERYRHIPPGEYVRFYGPCC 154
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 113 IOPEMHTTLPVFRISLPVLDPEVDSLTPWQHLVLDWMSSEYSHL-BDPFLKLSGFPCC 171
||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 155 WIFYHHLRQEIQLQEFTHDHFVREAAKFLRGLOVN--GSRPGTFVGHVHVRGDYVHVMPK 212
||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 172 WIFFHHLRQEIIRREFTLHDHLREDARQLSLGRIGPAGIRPRTYGVHVRGDYLEVMEN 231
||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 213 VVKGVVADRYLQOALDWPFRARYSSLIFFVVTSGNWCARENIDTSHGDVVFAGDGLGESP 272
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 232 RWKGVVGDRAVLQKAMDWFRARHKDPIFVVTSGNWCARENIDTSHGDVVFAGDGLGESP 291
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 273 AKDFALLTCNHTIMTIGTFGWAAYLTGSDTIYLANYTLPDSPFLKIFKPEAAFLPEWT 332
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 292 GKDFALLTCNHTIMTIGTFGWAAYLAGDVIYLANFILPDSEFLKIFKPEAAFLPEWV 351
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 333 GIAADLSPL 341
||| | | | | |
Db 352 GINADLSPL 360
||||| | | | | |

RESULT 7
FUT1_RABBIT
ID FUT1_RABBIT STANDARD; PRT; 373 AA.
AC Q10979;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-
D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)
DE (Fucosyltransferase 1).
GN FUT1 OR RFT-1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95238390; PubMed=7721792;
RA Hitoshi S., Kusunoki S., Kanazawa I., Tsuji S.;
RT "Molecular cloning and expression of two types of rabbit beta-
galactoside alpha 1,2-fucosyltransferase.";
RL J. Biol. Chem. 270:8844-8850(1995)
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
FORM IN TRANS CISTERNAE OF GOLGI.
CC -!- TISSUE SPECIFICITY: BRAIN.
CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF
MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X80226; CAA56513.1; -
CC InterPro: IPR002516; GT_11.
```



```

form the H blood group antigen."
Proc. Natl. Acad. Sci. U.S.A. 87:6674-6678(1990).
[2]
SEQUENCE FROM N.A., AND VARIANT ALLELES.
RP MEDLINE=97240210; PubMed=9122901;
RA Wagner F.F., Flegel W.A.;
RT "polymorphism of the h allele and the population frequency of
RT sporadic nonfunctional alleles."
RL Transfusion 37:284-290(1997).
[3]
VARIANT HIS-164.
RP MEDLINE=94286534; PubMed=7912436;
RA Kelly R.J., Ernst L.K., Larsen R.D., Bryant J.G., Robinson J.S.,
RA Lowe J.B.;
RT "Molecular basis for H blood group deficiency in Bombay (Oh) and
RT para-Bombay individuals."
RL Proc. Natl. Acad. Sci. U.S.A. 91:5843-5847(1994).
[4]
VARIANT ARG-242.
RP MEDLINE=97445117; PubMed=9299444;
RA Koda Y., Soejima M., Johnson P.H., Smart E., Kimura H.;
RT "Missense mutation of FUT1 and deletion of FUT2 are responsible for
RT Indian Bombay phenotype at ABO blood group system."
RL Biochem. Biophys. Res. Commun. 238:21-25(1997).
CC -!- FUNCTION: CREATES A SOLUBLE PRECURSOR OLIGOSACCHARIDE FUC-ALPHA
CC ((1,2)GALBETA-) CALLED THE H ANTIGEN WHICH IS AN ESSENTIAL
CC SUBSTRATE FOR THE FINAL STEP IN THE SOLUBLE A AND B ANTIGEN
CC SYNTHESIS PATHWAY. H AND SE ENZYME FUCOSYLATE THE SAME ACCEPTOR
CC SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -!- POLYMORPHISM: NONFUNCTIONAL MUTANT OF FUT1 ARE THE CAUSE OF THE H-
CC BOMBAY BLOOD GROUP.
CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
CC TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF
CC MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M35531; AAA52639.1; -.
CC EMBL; 269587; CAA93435.1; -.
CC PIR; A36047; A36047.
CC Genew; HGNC:4012; FUT1.
CC MIM; 211100; -.
CC Interpro; IPR002516; GT_11.
CC Pfam; PF01531; Glyco_transf_11; 1.
CC Signal-anchors; Glycosyltransferase; Polymorphism; Blood group antigen.
CC Signal-anchor; Golgi stack; Polymorphism; Blood group antigen.
CC DOMAIN 1 8 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 9 25 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
CC DOMAIN 26 365 LUMENAL, CATALYTIC (POTENTIAL).
CC FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT VARIANT 154 154 Y -> C (IN BOMBAY H-).
CC FT /FTid=VAR_003417.
CC FT L -> H (IN PARA-BOMBAY).
CC FT /FTid=VAR_009708.
CC FT W -> C (IN BOMBAY H-).
CC FT /FTid=VAR_003418.
CC FT L -> R (IN BOMBAY H-).
CC FT /FTid=VAR_009709.
CC FT V -> E (IN BOMBAY H-).

```

```

FT VARIANT 315 315 /FTid=VAR_003419.
FT A -> V (IN BOMBAY H-).
FT /FTid=VAR_003420.
FT V -> C (IN BOMBAY H-).
FT /FTid=VAR_003421.
SQ SEQUENCE 365 AA; 41251 MW; 4F4442EC375C9D9E CRC64;
Query Match 58.4%; Score 1080.5; DB 1; Length 365;
Best Local Similarity 59.7%; Pred. No. 2.8e-85;
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;
QY 12 MAHFILFVTVSTIFHWQQR-----LAKIQAMWELPVOIPVL-----ASTSK 53
DB 11 LAFLLVCVLSVIFFLHIHQDSFPHGLGLSDPDRILVTPPPVAIFCLPGTAMGPNASSC 70
QY 54 ALCPSQLRGMTTINAIKRLGNOMGEYATLYALAKMNGRAFIAPQAMHSTLAPIFRTILPV 113
DB 71 PQHPASLSGTWTVVPGNRTGNGMGQVATLLAQLNGRRAFILPAMHAALAPVFRITLPV 130
QY 114 LHSATASRIPTQNYHLNDWMEYRHIPGGEYVRFYPCSWTFYHHLRQELI-QEFTLHD 173
DB 131 LAPEVDSRTWPRELQLBDMSEYADL-RDPFLKLSGFPCSWTFYHHLRQELIREFTLHD 189
QY 174 HVREEAQKFLRGLOV--NGSRPCTFEVGVHVRGDDYVHMPKVKVGVVADRRYLLQCALDWF 231
DB 190 HLREEAQSVLQGLRTGDRPTFVGVHVRGDILOVMFQRMKGVVGDSSAYLRQANDWF 249
QY 232 RARYSSLIFFVVTSGMAWCRENIDTSHGDVVFAGDGGEGSPAKDFALLTQCNTHTMTIGT 291
DB 250 RARHEAPVFFVVTSGMWCRENIDTSGDVTFAGDGGQEAATPKWDFALLTQCNTHTMTIGT 309
QY 292 FGTHAAVLTGGDIYIANYLTLPDSPEFLKIFKPAFAFLPEWTGTAADLSPL 341
DB 310 FGTHAAVLTGGDIYIANYLTLPDSPEFLKIFKPAFAFLPEWTGTAADLSPL 359

```

```

RESULT 10
FUT1_PIG STANDARD; PRT; 365 AA.
AC Q29043; O19101;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-
DE D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)
DE (Fucosyltransferase 1).
GN FUT1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96217559; PubMed=8613146;
RA Cohnen S., Moutouris E., McKenzie I.F., Sandrin M.S.;
RT "Molecular cloning of the gene coding for pig alpha1-->2
RL fucosyltransferase."
RL Immunogenetics 44:76-79(1996).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS THR-103 AND GLN-286.
RX MEDLINE=97468270; PubMed=9321466;
RA Meijerink E., Fries R., Voegeli P., Masabanda J., Wigger G.,
RA Stricker C., Neuschwander S., Bertschinger H.O., Stranzinger G.;
RT "Two alpha(1,2) fucosyltransferase genes on porcine chromosome 6q11
RT are closely linked to the blood group inhibitor (S) and Escherichia
RT coli F18 receptor (ECF18R) loci."
RL Mamm. Genome 8:736-741(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RA Meijerink E., Neuschwander S., Fries R., Dinter A.,
RA Bertschinger H.U., Stranzinger G., Voegeli P.;
RI "Expression and activity of porcine alpha(1,2)fucosyltransferases

```



DB	283	ENTIPQDDLAYSHVSCDATALTPSPSTFGWMLGYLSKGAQVVYQDIISTNDNVNKKGVLPD	34
QY	324	EAAFLPEWTGIAAD 337 :   :	
DB	343	DDEFVPSWTSIMLD 356	
RESULT 13			
ID	FUT8_PIG	STANDARD;	PRT; 575 AA.
AC	P79282;		
DT	15-JUN-2002 (Rel. 41, Created)		
DT	15-JUN-2002 (Rel. 41, Last sequence update)		
DE	Alpha-(1,6)-fucosyltransferase (EC 2.4.1.68) (Glycoprotein 6-alpha-L-fucosyltransferase) (GDP-fucose--glycoprotein fucosyltransferase)		
DE	(GDP-L-Fuc-N-acetyl-beta-D-glucosaminide alphaL-6-fucosyltransferase)		
DE	(alpha1-6FucT) (Fucosyltransferase 8).		
GN	FUT8.		
SU	Sus scrofa (Pig).		
OC	Eumkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
OX	NCBI_TaxID=9823;		
[I]			
RN	SEQUENCE FROM N.A., SEQUENCE OF 102-130; 333-344; 402-415 AND		
RP	566-575, AND FUNCTIONION.		
RC	TISUE=Brain;		
RX	MEDLINE=97066976; PubMed=89103378;		
RA	Uozumi N., Yanagidani S., Miyoshi E., Ihara Y., Sakuma T., Gao C.-X., Teshima T., Fujii S., Shiba T., Taniguchi N.;		
RT	"Purification and cDNA cloning of porcine brain GDP-L-Fuc:N-acetyl-beta-D-glucosaminide alphaL-->6fucosyltransferase.";		
RL	J. Biol. Chem. 271:27810-27817(1996).		
-!	FUNCTION: Catalyzes the addition of fucose in alpha 1-6 linkage to the first GlcNAc residue, next to the peptide chains in N-glycans		
CC	The pH optimum is 7.		
CC	-! CATALYTIC ACTIVITY: GDP-L-fucose + N4-(N-acetyl-beta-D-glucosaminy		
CC	L-1,2-alpha-D-mannosyl-1,3-(R(1)-alpha-1,6)-beta-D-mannosyl-beta-N-acetyl-1,4-D-glucosaminy		
CC	L-1,4-N-acetyl-beta-D-glucosaminy = GP + N4-(N-acetyl-beta-D-glucosaminy		
CC	L-1,2-alpha-D-mannosyl-1,3-(R(1)-alpha-1,6)-beta-D-mannosyl-1,4-		
CC	beta-N-acetyl-D-glucosaminy L-1,4-(alpha-L-fucosyl-1,6)-N-acetyl-D-		
CC	glucosaminy]asparagine.		
CC	-! PATHWAY: Glycosylation.		
CC	-! SUBCELLULAR LOCATION: Type II membrane protein. Membrane-bound form in trans cisternae of Golgi (5y similarity).		
CC	-! TISSUE SPECIFICITY: HIGHEST EXPRESSION IN BRAIN.		
CC	-! SIMILARITY: CONTAINS 1 SH3 DOMAIN.		
CC	-! SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 23.		
CC	-----		
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstations at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in complete modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/annou or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; D86723; BAA13157.1;		
DR	InterPro; IPR001452; SH3.		
DR	Pfam; PF00018; SH3; 1.		
DR	SMART; SM00326; SH3; 1.		
DR	PROSITE; PS00002; SH3; 1.		
KW	Transferase; Glycosyltransferase; Transmembrane; Signal-anchor;		
KW	Golgi stack; SH3 domain.		
FT	DOMAIN 1 9		
FT	TRANSMEM 10 30		
FT	CYTOPLASMIC (POTENTIAL). SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN (POTENTIAL)). LUMENAL, CATALYTIC (POTENTIAL).		
FT	DOMAIN 31 575		
FT	DOMAIN 502 563		
FT	SITE 299 305		
FT	SITE 365 366		
FT	SH3-BINDING (POTENTIAL). IMPORTANT FOR DONOR SUBSTRATE BINDING (SIMILARITY).		



```
SQ SEQUENCE 575 AA; 66229 MW; 0F199D0BC2018F7B CRC64;
Query Match 5.08; Score 92; DB 1; Length 575;
Best Local Similarity 20.5%; Pred. No. 2.6;
Matches 60; Conservative 47; Mismatches 108; Indels 78; Gaps 16;
QY 14 HFLFVFTVSTIFHVQOQLAKIQAWEL-----PVQIPVLASTSKALGPSQLRGMWTIN 67
DB 225 HHVVCFMIA--YGTQRTLALESUNHRYATGWETVFPVSETCTDRSGSS--TGHWS-- 278
QY 68 AIGRLGNMGVEYATLYALAKNNGRPAFTPAQMHSTLAFIRITPLVLSATASRIWPQNY 127
DB 279 --GEVKDKNVQVVELPIVDSVHPRPYLPPLAVPEDLAD-----RLRVHGDDPA--VWVWSQ 330
QY 128 HL-----NDWREE-----YRHTPPGYVREFTGCPGCSWFFYHHLRQLQEFTH 172
DB 331 FYKYLIRPQPMLEKIEBATKLGFKHPVIGVHVRTDKVGAEAFAHP-----IEEYIV- 384
QY 173 DHVREACKFLRGLQVNGSR-----PGTFVGHVRRGDYVHVMPK--VW----- 214
DB 365 -HVEDFQLLARRMOVDKRYLATDDPALLKEAKTKYPSYEFISDINSAGLHNRYT 443
QY 215 -----KGVVADRYLQOALDFWARYSSLIFFVTSNGMAWCR---ENIDTSHGD 260
DB 444 ENSLRGVILDIHFLSQA-DFLVCTFSSQV-----CRVAYEIMOALHPD 485
RESULT 14
ID FUT8_HUMAN STANDARD; PRT; 575 AA.
AC Q9BYC5; Q9BYC6; Q9P205; Q9P206; Q00235;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha-(1,6)-fucosyltransferase (EC 2.4.1.68) (Glycoprotein 6-alpha-L-
DE fucosyltransferase) (GDP-fucose-glycoprotein fucosyltransferase)
DE (GDP-L-Fuc:N-acetyl-beta-D-glucosaminide alpha1,6-fucosyltransferase)
DE (alpha1-6FucT) (Fucosyltransferase 8).
GN FUT8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND FUNCTION.
RX MEDLINE=97279058; PubMed=9133635;
RA Yanagidani S., Uozumi N., Ihara Y., Miyoshi E., Yamaguchi N.,
RA Taniguchi N.;
RT "Purification and cDNA cloning of GDP-L-Fuc:N-acetyl-beta-D-
RT glucosaminide:alpha1-6 fucosyltransferase (alpha1-6 FucT) from human
RT gastric cancer MKN45 cells.";
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Embryo;
RA Caillaud A., Balanzino L., Candellier J.J., Oriol R., Mollicone R.;
RT "Differential splice variants of human FUT8 embryonic cDNA.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=20275614; PubMed=10814706;
RA Yamaguchi Y., Ikeda Y., Takahashi T., Ihara H., Tanaka T., Sasho C.,
RA Uozumi N., Yanagidani S., Inoue S., Fujii J., Taniguchi N.;
RT "Genomic structure and promoter analysis of the human
RT alpha1,6-fucosyltransferase gene (FUT8).";
RL Glycobiology 10:637-643(2000).
RN [4]
RP DONOR SUBSTRATE BINDING, AND MUTAGENESIS OF ARG-365 AND ARG-366.
RX MEDLINE=20229550; PubMed=10764839;
RA Takahashi T., Ikeda Y., Tateishi A., Yamaguchi Y., Ishikawa M.,
RA Taniguchi N.;
RT "A sequence motif involved in the donor substrate binding by
RT alpha1,6-fucosyltransferase: the role of the conserved arginine
residues.";
RL Glycobiology 10:503-510(2000).
-!- FUNCTION: Catalyzes the addition of fucose in alpha 1-6 linkage to
the first GlcNAc residue, next to the peptide chains in N-glycans.
-!- CATALYTIC ACTIVITY: GDP-L-fucose + N4-(N-acetyl-beta-D-
glucosaminyl)-1,2-alpha-D-mannosyl-1,3-(R(1)-alpha-1,6)-beta-D-
mannosyl-beta-N-acetyl-1,4-D-glucosaminyl-1,4-N-acetyl-D-
glucosaminyl)asparagine = GDP + N4-(N-acetyl-beta-D-glucosaminyl-
1,2-alpha-D-mannosyl-1,3-(R(1)-alpha-1,6)-beta-D-mannosyl-1,4-
beta-N-acetyl-D-glucosaminyl-1,4-(alpha-L-fucosyl-1,6)-N-acetyl-D-
glucosaminyl)asparagine.
-!- PATHWAY: Glycosylation.
-!- SUBCELLULAR LOCATION: Type II membrane protein. Membrane-bound
form in trans cisternae of Golgi (By similarity).
-!- ALTERNATIVE PRODUCTS: 2 isoforms: 1 (shown here) and 2/retinal;
are produced by alternative splicing.
-!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
-!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 23.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; D89289; BAA19764.1; -
EMBL; AF038281; AB92372.2; -
EMBL; AF038280; AB92372.2; JOINED.
EMBL; Y17979; CAA76988.1; -
EMBL; Y17978; CAA76985.1; -
EMBL; Y17977; CAA76986.1; -
EMBL; Y17978; CAA76987.1; -
EMBL; AB049828; BAB40975.1; -
EMBL; AB049740; BAB40929.2; -
EMBL; AB032573; BAA92859.2; -
EMBL; AB032572; BAA92858.1; -
EMBL; AB032568; BAA92858.1; JOINED.
EMBL; AB032569; BAA92858.1; JOINED.
EMBL; AB032570; BAA92858.1; JOINED.
EMBL; AB032571; BAA92858.1; JOINED.
Genew; HGNC:4019; FUT8.
MIM; 602589; -
DR InterPro; IPR001452; SH3.
DR Pfam; PF00018; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50002; SH3; FALSE_NEG.
KW Transferase; Glycosyltransferase; Transmembrane; Signal-anchor;
KW Golgi_stack; SH3_domain; Alternative_splicing.
FT DOMAIN 1 9
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT TRANSMEM 10 30
FT POTENTIAL.
FT LUMENAL, CATALYTIC (POTENTIAL).
FT SH3.
FT SH3-BINDING (POTENTIAL).
FT IMPORTANT FOR DONOR SUBSTRATE BINDING.
FT EVKDKNVQVVELPIVDSVHPRPYLPPLAVPEDLADRLVRVH
FT GDPVWWS -> TETMLLVITLPGQLDCTIDTQKIHVF
FT E (IN ISOFORM 2).
FT MISSING (IN ISOFORM 2).
FT R->A, K: COMPLETE LOSS OF ACTIVITY.
FT R->A, K: DECREASES ACTIVITY TO 3%.
FT SEQUENCE 575 AA; 66515 MW; 5AE24A93881E18D0 CRC64;
Query Match 4.9%; Score 90; DB 1; Length 575;
Best Local Similarity 20.2%; Pred. No. 3.8;
Matches 59; Conservative 46; Mismatches 111; Indels 76; Gaps 15;
QY 14 HFLFVFTVTS-----TIFHVQOQLAKIQAWELFQIPVLASTSKALGPSQLRGMWTINA 68
DB 225 HHVVCFMIA YGTQRTLALESUNHRYATGWETVFPVSETCTDRSGIS--TGHWS--- 278
```

```

DR Pfam; PF000018; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00002; SH3; FALSE_NEG.
KW Transferase; Glycosyltransferase; Transmembrane; Signal-anchor;
KW Golgi stack; SH3 domain.
FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 10 30 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT 31 575 LUMENAL, CATALYTIC (POTENTIAL).
FT DOMAIN 502 563 SH3.
FT SITE 299 305 SH3-BINDING (POTENTIAL).
FT SITE 365 366 IMPORTANT FOR DONOR-SUBSTRATE BINDING.
FT CONFLICT 40 40 T -> S (IN REF. 2).
FT CONFLICT 388 388 Q -> E (IN REF. 2).
FT CONFLICT 418 418 N -> K (IN REF. 2).
SQ SEQUENCE 575 AA; 66555 MW; 7BE2ED3146E0C45F CRC64;

Query Match 4.9%; Score 90; DB 1; Length 575;
Best Local Similarity 19.8%; Pred. No. 3.8;
Matches 58; Conservative 49; Mismatches 108; Indels 78; Gaps

QY 14 HPILEFVTVS-----TIFHVQQLAKIQAMWELPVOIPVLASTSKALGPSQLRGWTTINA 68
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 225 HHVYCFMIAVGTQTILLESQNWRYATGGWE-TVPRPYSETCTDRSGLS--TGHSW--- 278

QY 69 IGRIGNOMGEYATILYALAKNGRGPATIPQMHSTLAP-IPRTPLVHSTATSRIPQWY 127
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 279 -GEYNDKNIQVSLPVDLSLHPRPPLPLAVPDLADRLRV-----HGDPA--VVMV 330

QY 128 HL-----NDWMEEE-----YRHIPPGEVVRITGVPSCWTFVHHLRQKILQEFTLH 172
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 331 FVKYILRPQPLWKEETAEATKKLOFKHPVIGVHVRITDKVGTAEACHPEEYMW----- 384

QY 173 DHVREEACKPLRGIVNGSR-----PGTFVGHVHRRGDYHVMPK--VW----- 214
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 385 -HYEQHQLLARQMDKKEVYLTDPTLLKEANKYSNYEFLSDNSISWSAGLHRRYT 443

QY 215 ----KGVVADRRYIQQALDWFRARYSSLIPVVTNSGMWACR-----ENIDTSHGD 260
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 444 ENSLRGVIIDHLSQA-DELVCVFSSV-----CRVAYEKMTLHPD 485

```

```

AN SEQUENCE FROM N.A.
AR Submitted R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes the addition of fucose in alpha 1-6 linkage to
CC the first GlcNAc residue, next to the peptide chains in N-glycans
CC (By similarity).
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + N4-(N-acetyl-beta-D-
CC glucosaminyl-1,2-alpha-D-mannosyl-1,3-(R(1)-alpha-1,6)-beta-D-
CC mannosyl-beta-N-acetyl-1,4-D-glucosaminyl-1,4-N-acetyl-D-
CC glucosaminyl)asparagine = GDP + N4-(N-acetyl-beta-D-glucosaminyl-
CC 1,2-alpha-D-mannosyl-1,3-(R(1)-alpha-1,6)-beta-D-mannosyl-1,4-
CC beta-N-acetyl-D-glucosaminyl-1,4-(alpha-1-fucosyl-1,6)-N-acetyl-D-
CC glucosaminyl)asparagine.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Membrane-bound
CC form in trans cisternae of Golgi (By similarity).
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 23.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collabora-
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no
CC modified and this statement is not removed. Usage by and for commer-
CC entities requires a license agreement (See http://www.isb-sib.ch/announc
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; ABC025198; BAA76392.1; -
CC EMBL; BC010666; RAH10666.1; -
CC EMBL; MGI:1858901; Fut8
CC InterPro: IPR001452; SH3.

```

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:01:52 ; Search time 36.6774 Seconds  
(without alignments)  
1932.533 Million cell updates/sec

Title: US-10-040-863-11

Perfect score: 1850

Sequence: 1 MLVQMPSPFMAHFLVF.....AAFLPEWTGTAADLSPLKH 344

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1825.5	98.7	343	6	Q9TUD3
2	1821.5	98.5	343	6	Q77486
3	1815.5	98.1	343	6	Q77485
4	1793.5	96.9	338	6	Q9TUD5
5	1784.5	96.5	343	6	Q9TTC7
6	1771.5	95.8	332	4	Q99450
7	1766.5	95.5	343	6	Q77487
8	1736.5	93.9	343	6	Q77412
9	1566.5	84.7	346	4	Q14338
10	1521.5	82.2	347	6	Q29505
11	1483	80.2	354	11	O35087
12	1481	80.1	354	11	Q35087
13	1475.5	79.8	347	11	Q9J275
14	1475	79.7	380	11	Q9J227
15	1471.5	79.5	347	11	Q9J444
16	1409.5	76.2	273	6	Q9TSL5

17	1376	74.4	353	11	O55025
18	1369.5	74.0	321	11	Q91VF0
19	1368.5	74.0	321	11	Q920W2
20	1354.5	73.2	321	11	Q920W3
21	1344.5	72.7	273	6	Q9TTC9
22	1335.5	72.2	273	6	Q9TUD0
23	1332.5	72.0	348	6	Q9TUD1
24	1311.5	70.9	332	6	Q9N266
25	1304.5	70.5	368	11	Q920T3
26	1301.5	70.4	368	11	Q920T4
27	1298.5	70.2	368	11	Q91V73
28	1295.5	70.0	368	11	Q920T2
29	1293.5	69.9	368	6	Q9TTC3
30	1233.5	66.7	299	11	Q9WUE6
31	1100.5	59.5	360	6	Q9TTC7
32	1092	59.0	377	11	Q9JL28
33	1092	59.0	377	11	Q9J327
34	1084	58.6	366	6	Q9TUD4
35	1082	58.5	377	11	Q920W5
36	1082	58.5	377	11	Q920W4
37	1080.5	58.4	365	6	Q9TUE7
38	1076.5	58.2	365	4	Q8WXB6
39	1076	58.2	366	6	Q9TUE6
40	1075.5	58.1	365	4	O14503
41	1072	57.9	366	6	Q9TUD6
42	1071.5	57.9	365	4	O14506
43	1071.5	57.9	365	4	O14507
44	1071	57.9	366	6	Q9TUD2
45	1070.5	57.9	353	6	Q9N270

#### ALIGNMENTS

RESULT 1

Q9TUD3

ID Q9TUD3 PRELIMINARY; PRT; 343 AA.

AC Q9TUD3; Q9TUD3; 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Alpha (1,2) fucosyl transferase.

GN FUT2

OS Gorilla gorilla (gorilla).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.

OX NCBI\_TaxID=9593;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ALEXIS;

RX MEDLINE=20188794; PubMed=10723735;

RA Apoll P.-A., Roubinet F., Despiau S., Mollicone R., Oriol R.,

RA Blancher A.;

RT "Evolution of alpha2-Fucosyltransferase genes in primates: relation between an intronic Alu-Y element and red cell expression of ABH antigens.";

RL Mol. Biol. Evol. 17:337-351(2000).

DR EMBL; AF080606; AAF14068.1; -

DR InterPro; IPR002516; GT\_11.

DR Pfam; PF01531; Glyco\_transf\_11; 1.

KW Transferase.

SK SEQUENCE 343 AA; 39001 MW; A753375D47AE8C8 CRC64;

Query Match 98.7%; Score 1825.5; DB 6; Length 343;

Best Local Similarity 99.4%; Pred. No. 6.7e-162;

Matches 342; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 MLVQMPSPFMAHFLVFVTSTFHVQQRQAKIQAMWELPVQIPVLASTSKALGPSQL 60

Db 1 MLVQMPSPFMAHFLVFVTSTFHVQQRQAKIQAMWELPVQIPVLASTSKALGPSQL 60

Qy 61 RGMWTINAIKRLGNMGYATLYALAKNGRPAPFPAQMHSHTLAFIFRITLPVLHSATAS 120

```

Db 61 RGMWTINAIGRLGNQMGGEATLYALAKMNGRPAFIPQAMHSTLAPIFRITLPLVLSATAS 120
QY 121 RIPQWYHLNDWMEEEYRHIPPGEYVFTGTCSTWTFYHHLRQELQFTLHDHVRREAQ 180
DB 121 RIPQWYHLNDWMEEEYRHIPPGEYVFTGTCSTWTFYHHLRQELQFTLHDHVRREAQ 179
QY 181 KFLRGLQVNSRPGTFVGVHVRGDDYVHVMKVKWGVVADRRYLOQALDWFRRYSSLIIF 240
DB 180 KFLRGLQVNSRPGTFVGVHVRGDDYVHVMKVKWGVVADRRYLOQALDWFRRYSSPIF 239
QY 241 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLT 300
DB 240 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLT 299
QY 301 GGDITLYANITLDPSPFLKIKPPEAFLPEWTGIAADLSPLK 344
DB 300 GGDITLYANITLDPSPFLKIKPPEAFLPEWTGIAADLSPLK 343

RESULT 2
O77486 PRELIMINARY; PRT: 343 AA.
AC O77486;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Alpha(1,2) fucosyltransferase.
GN FUT2.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;
RT "The old origin of a null allele se428 of the human ABO-secretor type
RT alpha(1,2) fucosyltransferase gene (FUT2).";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennoh G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human secretor blood
RT group alpha(1,2) fucosyltransferase gene (FUT2).";
RL J. Biol. Chem. 270:4640-4649(1995).
DR EMBL; AB015635; BAA31128.1;
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 343 AA; 38973 MW; 894E28BD74AE8FEC CRC64;

Query Match 98.5%; Score 1821.5; DB 6; Length 343;
Best Local Similarity 99.1%; Pred. No. 1.6e-161;
Matches 341; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 MLVVQMPFSPMAHFIILFVFTVSTIFHVQORLAKIOAMWELPQVPIVLASTSKALGPSQL 60
DB 1 MLVVQMPFSPMAHFIILFVFTVSTIFHVQORLAKIOAMWELPQVPIVLASTSKALGPSQL 60
QY 61 RGMWTINAIGRLGNQMGGEATLYALAKMNGRPAFIPQAMHSTLAPIFRITLPLVLSATAS 120
DB 61 RGMWTINAIGRLGNQMGGEATLYALAKMNGRPAFIPQAMHSTLAPIFRITLPLVLSATAS 120
QY 121 RIPQWYHLNDWMEEEYRHIPPGEYVFTGTCSTWTFYHHLRQELQFTLHDHVRREAQ 180
DB 121 RIPQWYHLNDWMEEEYRHIPPGEYVFTGTCSTWTFYHHLRQELQFTLHDHVRREAQ 179
QY 181 KFLRGLQVNSRPGTFVGVHVRGDDYVHVMKVKWGVVADRRYLOQALDWFRRYSSLIIF 240
DB 180 KFLRGLQVNSRPGTFVGVHVRGDDYVHVMKVKWGVVADRRYLOQALDWFRRYSSPIF 239
QY 241 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLT 300
DB 240 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLT 299

```

```

Db 240 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLT 299
QY 301 GGDITLYANITLDPSPFLKIKPPEAFLPEWTGIAADLSPLK 344
DB 300 GGDITLYANITLDPSPFLKIKPPEAFLPEWTGIAADLSPLK 343

RESULT 3
O77485 PRELIMINARY; PRT: 343 AA.
AC O77485;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Alpha(1,2) fucosyltransferase.
GN FUT2.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;
RT "The old origin of a null allele se428 of the human ABO-secretor type
RT alpha(1,2) fucosyltransferase gene (FUT2).";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennoh G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human secretor blood
RT group alpha(1,2) fucosyltransferase gene (FUT2).";
RL J. Biol. Chem. 270:4640-4649(1995).
DR EMBL; AB015634; BAA31127.1;
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 343 AA; 38999 MW; B849D7E152852081 CRC64;

Query Match 98.1%; Score 1815.5; DB 6; Length 343;
Best Local Similarity 98.8%; Pred. No. 5.8e-161;
Matches 340; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 MLVVQMPFSPMAHFIILFVFTVSTIFHVQORLAKIOAMWELPQVPIVLASTSKALGPSQL 60
DB 1 MLVVQMPFSPMAHFIILFVFTVSTIFHVQORLAKIOAMWELPQVPIVLASTSKALGPSQL 60
QY 61 RGMWTINAIGRLGNQMGGEATLYALAKMNGRPAFIPQAMHSTLAPIFRITLPLVLSATAS 120
DB 61 RGMWTINAIGRLGNQMGGEATLYALAKMNGRPAFIPQAMHSTLAPIFRITLPLVLSATAS 120
QY 121 RIPQWYHLNDWMEEEYRHIPPGEYVFTGTCSTWTFYHHLRQELQFTLHDHVRREAQ 180
DB 121 RIPQWYHLNDWMEEEYRHIPPGEYVFTGTCSTWTFYHHLRQELQFTLHDHVRREAQ 179
QY 181 KFLRGLQVNSRPGTFVGVHVRGDDYVHVMKVKWGVVADRRYLOQALDWFRRYSSLIIF 240
DB 180 KFLRGLQVNSRPGTFVGVHVRGDDYVHVMKVKWGVVADRRYLOQALDWFRRYSSPIF 239
QY 241 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLT 300
DB 240 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLT 299
QY 301 GGDITLYANITLDPSPFLKIKPPEAFLPEWTGIAADLSPLK 344
DB 300 GGDITLYANITLDPSPFLKIKPPEAFLPEWTGIAADLSPLK 343

RESULT 4
Q9TUD5 PRELIMINARY; PRT: 338 AA.
AC Q9TUD5;
DT 01-MAY-2000 (Tremblrel. 13, Created)

```

RT between an intronic Alu-Y element and red cell expression of ABH  
 RT antigens.";  
 RL Mol. Biol. Evol. 17:337-351(2000).  
 DR EMBL: AF136648; AAF25585.1; -.  
 DR InterPro: IPR002516; GT\_11.  
 DR Pfam: PF01531; Glyco\_transf\_11; 1.  
 DR Glycosyltransferase; Transferase.  
 KW NCBI\_TaxID=9598;  
 SQ SEQUENCE 343 AA: 39059 MW: A52161BE67E1F64D CRC64;  
 Query Match 96.5%; Score 1784.5; DB 6; Length 343;  
 Best Local Similarity 97.4%; Pred. No. 4.5e-158;  
 Matches 335; Conservative 2; Mismatches 6; Indels 1; Gaps 1;  
 QY 1 MLVVQMPFPPMAHFLFVFTVSTIFHVQORLAKIOAMWELPVQIPVLASTSKALGPSQL 60  
 Db 1 MLVVQMPFPPMAHFLFVFTVSTIFHVQORLAKIOAMWELPVQIPVLASTSKALGPSQL 60  
 QY 61 RGMWTINAGRLGNQGEYATLYALAKMNGRPAFIPAQMHSSTIAPFRITLPVLHSATAS 120  
 Db 61 RGMWTINAGRLGNQGEYATLYALAKMNGRPAFIPAQMHSSTIAPFRITLPVLNSAMAS 120  
 QY 121 RIPWQNYHLNDWMEEEYRHIPGCVYVRFYHHLRQEILOEFTLHDHVRDEAQ 180  
 Db 121 RIPWQNYHLNDWMEEEYRHIPGCVYVRFYHHLRQEILOEFTLHDHVRDEAQ 179  
 QY 181 KFLRGLOVNGSRGTGTVGVVIVRGGDYVHVMPKVGVDVRRYLOQALDWFERYSSLIF 240  
 Db 180 KFLRGLOVNGSRGTGTVGVVIVRGGDYVHVMPKVGVDVRRYLOQALDWFERYSSPIF 239  
 QY 241 VVTSNGMWCRENIDTSHGDVTFAGDIEGSPAKDFALLTQCNTIMTIGTFGIWAAYLT 300  
 Db 240 VVTSNGMWCRENIDTSHGDVTFAGDIEGSPAKDFALLTQCNTIMTIGTFGIWAAYLT 299  
 QY 301 GGDFTYLANLYTLDPSPFLKIFKPEAAFLPEWTGTAADLSPLLK 344  
 Db 300 GGDFTYLANLYTLDPSPFLKIFKPEAAFLPEWTGTAADLSPLLK 343  
 RESULT 6  
 Q99450 PRELIMINARY; PRT; 332 AA.  
 AC Q99450;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Alpha (1,2) fucosyltransferase.  
 GN SEJ.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96199252; PubMed=8621666;  
 RA Kudo T., Iwasaki H., Shinya N., Nishihara S., Ando T., Narimatsu I.,  
 RA Narimatsu H.;  
 RA "Molecular Genetic Analysis of Human Histo-blood Group System. II.  
 RA Secretor Gene Inactivation by a Novel Single Missense Mutation A365T  
 RA in Japanese Nonsecretor Individuals.";  
 RL J. Biol. Chem. 271:9830-9837(1996).  
 DR EMBL: D89326; BAA13543.1; -.  
 DR InterPro: IPR002516; GT\_11.  
 DR Pfam: PF01531; Glyco\_transf\_11; 1.  
 DR Glycosyltransferase; Transferase.  
 KW NCBI\_TaxID=9580;  
 SQ SEQUENCE 332 AA: 37773 MW: 092A4C317DA45986 CRC64;  
 Query Match 95.8%; Score 1771.5; DB 4; Length 332;  
 Best Local Similarity 99.4%; Pred. No. 7e-157;  
 Matches 331; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 12 MAHFILFVFTVSTIFHVQORLAKIOAMWELPVQIPVLASTSKALGPSQLRGMTNAIGR 71  
 Db 1 MAHFILFVFTVSTIFHVQORLAKIOAMWELPVQIPVLASTSKALGPSQLRGMTNAIGR 60

RT between an intronic Alu-Y element and red cell expression of ABH  
 RT antigens.";  
 RL Mol. Biol. Evol. 17:337-351(2000).  
 DR EMBL: AF080604; AAF14066.1; -.  
 DR InterPro: IPR002516; GT\_11.  
 DR Pfam: PF01531; Glyco\_transf\_11; 1.  
 DR Transferase.  
 KW NCBI\_TaxID=9598;  
 SQ SEQUENCE 338 AA: 38428 MW: E7FABA0FF1BC95F9 CRC64;  
 Query Match 96.9%; Score 1793.5; DB 6; Length 338;  
 Best Local Similarity 98.8%; Pred. No. 6.4e-159;  
 Matches 335; Conservative 1; Mismatches 2; Indels 1; Gaps 1;  
 QY 6 MPFSFMAHFLFVFTVSTIFHVQORLAKIOAMWELPVQIPVLASTSKALGPSQLRGMT 65  
 Db 1 MPFSFMAHFLFVFTVSTIFHVQORLAKIOAMWELPVQIPVLASTSKALGPSQLRGMT 60  
 QY 66 INALGRGNQGEYATLYALAKMNGRPAFIPAQMHSSTIAPFRITLPVLHSATASRIPWQ 125  
 Db 61 INALGRGNQGEYATLYALAKMNGRPAFIPAQMHSSTIAPFRITLPVLHSATASRIPWQ 120  
 QY 126 NYHLNDWMEEEYRHIPGCVYVRFYHHLRQEILOEFTLHDHVRDEAQKFLRG 185  
 Db 121 NYHLNDWMEEEYRHIPGCVYVRFYHHLRQEILOEFTLHDHVRDEAQKFLRG 179  
 QY 186 LOVNGSRPGTVGVVIVRGGDYVHVMPKVGVDVRRYLOQALDWFERYSSLIFVVTSN 245  
 Db 180 LOVNGSRPGTVGVVIVRGGDYVHVMPKVGVDVRRYLOQALDWFERYSSPIFVVTSN 239  
 QY 246 GMWCRENIDTSHGDVTFAGDIEGSPAKDFALLTQCNTIMTIGTFGIWAAYLTGGDTI 305  
 Db 240 GMWCRENIDTSHGDVTFAGDIEGSPAKDFALLTQCNTIMTIGTFGIWAAYLTGGDTI 299  
 QY 306 YLANLYTLDPSPFLKIFKPEAAFLPEWTGTAADLSPLLK 344  
 Db 300 YLANLYTLDPSPFLKIFKPEAAFLPEWTGTAADLSPLLK 338  
 RESULT 5  
 Q99450 PRELIMINARY; PRT; 343 AA.  
 AC Q99450;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Alpha(1,2)fucosyltransferase.  
 GN FUT2.  
 OS Hylobates lar (Common gibbon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.  
 OX NCBI\_TaxID=9580;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=TJGA;  
 RX MEDLINE=20186794; PubMed=10723735;  
 RA Apoll P.-A., Roubinet F., Desplau S., Mollicone R., Oriol R.,  
 RA Blancher A.;  
 RA "Evolution of alpha2-Fucosyltransferase genes in primates: relation



Q14338  
ID Q14338 PRELIMINARY; PRT; 346 AA.  
AC Q14338;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Alpha (1, 2) fucosyltransferase.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_taxid=9606;  
RN [1]  
RP PARTIAL SEQUENCE FROM N.A.  
RA Kimura H.;  
RL Submitted (JAN-1996) to the EMBL/GenBank/DDJB databases.  
RN [2]  
RP SEQUENCE OF 84-106 FROM N.A.  
RX MEDLINE=96335703; PubMed=8755920;  
RA Koda Y., Soejima M., Liu Y., Kimura H.;  
RT "Molecular basis for secretor type alpha(1,2)-fucosyltransferase gene  
deficiency in a Japanese population: a fusion gene generated by  
unequal crossover responsible for the enzyme deficiency.";  
RL Am. J. Hum. Genet. 59:343-350(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95181460; PubMed=7876235;  
RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;  
RT "Sequence and expression of a candidate for the human secretor blood  
group alpha(1,2) fucosyltransferase gene (FUT2).";  
RL J. Biol. Chem. 270:4640-4649(1995).  
DR EMBL; D82933; BAAL1638.1; -;  
DR InterPro; IPR002516; GT\_11.  
DR Pfam; PF01531; Glyco.transf.11; 1.  
KW Glycosyltransferase; Transferrase.  
SQ SEQUENCE 346 AA; 39395 MW; 5D72275F8571410B CRC64;

Query Match 84.7%; Score 1566.5; DB 4; Length 346;  
Best Local Similarity 87.3%; Pred. No. 9.7e-138;  
Matches 295; Conservative 9; Mismatches 23; Indels 11; Gaps 3;

QY 9 SFPMAHFLVFFVVSIFHVOORLAKIQAMWELPQVLPVLASTKALGPSOL--RGMWII 66  
DB 18 SFTVFVFVFAIFVVSIFHVOORLAKIQAMWELPQVLPVLASTKALGPSOL--RGMWII 69

QY 67 NAIQRLGNQMGVEATLYALAKMNGRPAPFPAQMHSITLAPFRITLPVLSATASRIPWQN 126  
DB 70 NSKGRNGQMGVEATLYALAKMNGRPAPFPAQMHSITLAPFRITLPVLSATASRIPWQN 129

QY 127 YHLNDWMEERYRHIPPGVEYVFRTGYPCSWTFYHHLROEILQEFTHDHRVREAAKFLRGL 186  
DB 130 YHLNDWMEERYRHIPPGVEYVFRTGYPCSWTFYHHLROEILQEFTHDHRVREAAKFLRGL 188

QY 187 QVNSRSGTGVGVHVRGDDYVHVMPKWKGVVADRRYLQALDWFRRYSLLIFVWTSNG 246  
DB 189 QVNSRSGTGVGVHVRGDDYVHVMPKWKGVVADRRYLQALDWFRRYSLLIFVWTSNG 248

QY 247 MAWCRENIDTSHGDVVFAGDGIKSPAKDFALLTQCNTIMTIGTFGIWAAYLTGGDTIY 306  
DB 249 MAWCRENIDTSHGDVVFAGDGIKSPAKDFALLTQCNTIMTIGTFGIWAAYLTGGDTIY 308

QY 307 LANTLPDPSPLKTFKPEAFLPEWTGTAADLSPLKKH 344  
DB 309 LANTLPDPSPLKTFKPEAFLPEWTGTAADLSPLKKH 346

RESULT 10  
Q29505  
ID Q29505 PRELIMINARY; PRT; 347 AA.  
AC Q29505;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Galactoside 2-L-fucosyltransferase 3 (EC 2.4.1.69) (Secretor blood

DE group ALPHA-2-fucosyltransferase) (GDP-L-fucose:beta-D-galactoside  
DE 2-ALPHA-L-fucosyltransferase) (ALPHA(1,2)F) (Fucosyltransferase 3).  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_taxid=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=GASTROINTESTINAL TRACT;  
RX MEDLINE=96279281; PubMed=8663168;  
RA Hitoshi S., Kojima N., Kanazawa I., Tsuji S.;  
RT "Molecular cloning and expression of a third type of rabbit GDP-L-  
fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase.";  
RL J. Biol. Chem. 271:16975-16981(1996).  
CC -!- FUNCTION: CREATES A MEMBRANE-ASSOCIATED PRECURSOR OLIGOSACCHARIDE  
FUC-ALPHA(1,2)GALBETA-) CALLED THE H ANTIGEN WHICH IS AN  
ESSENTIAL SUBSTRATE FOR THE FINAL STEP IN THE MEMBRANE-ASSOCIATED  
A AND B ANTIGEN SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE THE  
SAME ACCEPTOR SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.  
CC -!- CATALYTIC ACTIVITY: GDP-L-FUCOSE + BETA-D-GALACTOSYL-R = GDP +  
ALPHA-L-FUCOSYL-1,2-BETA-D-GALACTOSYL-R.  
CC -!- PATHWAY: GLYCOSYLATION.  
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND  
FORM IN TRANS CISTERNAE OF GOLGI.  
CC -!- TISSUE SPECIFICITY: ADULT GASTROINTESTINAL TRACT AND COLON.  
CC -!- MISCELLANEOUS: THERE ARE THREE GENES (FUT1, FUT2 AND FUT3) WHICH  
ENCODE GALACTOSIDE 2-L-FUCOSYLTRANSFERASE IN RABBIT. THEY ARE  
EXPRESSED IN A TISSUE-SPECIFIC MANNER.  
CC -!- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN  
GLYCOSYLTRANSFERASES.  
DR EMBL; X91269; CAA62669.1; -;  
DR InterPro; IPR002516; GT\_11.  
DR Pfam; PF01531; Glyco.transf.11; 1.  
KW Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane;  
KW Signal-anchor; Golgi stack.  
FT DOMAIN 1 5 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 6 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
(POTENTIAL).  
FT DOMAIN 29 347 LUMENAL, CATALYTIC (POTENTIAL).  
FT CARBOHYD 192 192 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 286 286 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 312 312 N-LINKED (GLCNAC...) (POTENTIAL).  
SQ SEQUENCE 347 AA; 39469 MW; 746F7007309862A5 CRC64;

Query Match 82.2%; Score 1521.5; DB 6; Length 347;  
Best Local Similarity 81.6%; Pred. No. 1.9e-133;  
Matches 284; Conservative 19; Mismatches 40; Indels 5; Gaps 3;

QY 1 MLVVOMPSPRPMAHFLVFTVSTIFHVOORLAKIQAMWE--LPVQIPVLA--STSKALG 56  
DB 1 MSTAQVFPFAPVWVHVLVFTASTIFHVOORLAKIQAMWE--LPVQIPVLA--STSKALG 59

QY 57 PSQLRGWTINAIQRLGNQMGVEATLYALAKMNGRPAPFPAQMHSITLAPFRITLPVLSH 116  
DB 61 SRPLGGMTINAMGRNGQMGVEATLYALAKMNGRPAPFPAQMHSITLAPFRITLPVLSH 120

QY 117 ATASRIPWQNTHLNDWMEERYRHIPPGVEYVFRTGYPCSWTFYHHLROEILQEFTHDHRV 176  
DB 121 STASRVPWQNTHLNDWMEERYRHIPPGVEYVFRTGYPCSWTFYHHLROEILQEFTHDHRV 179

QY 177 EEAQFLRGLGVNSRSGTGVGVHVRGDDYVHVMPKWKGVVADRRYLQALDWFRRYS 236  
DB 180 EEAQFLRGLRVNSRSGTGVGVHVRGDDYVHVMPKWKGVVADRRYLQALDWFRRYS 239

QY 237 SLIFVWTSNGMAWCRENIDTSHGDVVFAGDGIKSPAKDFALLTQCNTIMTIGTFGIWA 296  
DB 240 SPVFWVTSNGMAWCRENIDTSHGDVVFAGDGIKSPAKDFALLTQCNTIMTIGTFGIWA 299

QY 297 AYLTGGDTIYIANYTLPDPSPLKTFKPEAFLPEWTGTAADLSPLKKH 344  
DB 300 AYLTGGDTIYIANYTLPDPSPLKTFKPEAFLPEWTGTAADLSPLKKH 347

```
RESULT 11
Q35087 PRELIMINARY; PRT; 354 AA.
AC O35087;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase.
GN FTB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-COLON CANCER;
RA Seojima M., Wang B., Koda Y., Kimura H.;
RT "Two distinct rat GDP-L-fucose:b-D-galactoside 2-a-L-
RT fucosyltransferase genes.";
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB006138; BAA21742.1; -.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF04531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 354 AA; 39995 MW; 8636444889215BA1 CRC64;

Query Match 80.2%; Score 1493; DB 11; Length 354;
Best Local Similarity 77.7%; Pred. No. 6.1e-130;
Matches 276; Conservative 30; Mismatches 37; Indels 12; Gaps 3;

QY 1 MLVQMPFPMAHFLFVFTVSTIFHVQORLAKIQAM--WELPV-----QIPVLA 49
DB 1 MASAQVPFSEPLAHFLFVFTVSTIHLQORIVKLOPLSEKELPMITQSSSGNTESPEMR 60
QY 50 STSKALGPSOLRGWMTINAGRLGNQMGVEYATLYALAKMNGRPAFIPAQMHSILAFIRI 109
DB 61 RDSEQHNGELRGWMTINAGRLGNQMGVEYATLYALAKMNGRPAFIPAQMHSILAFIRI 120
QY 110 TLPVLSATASRIIPWQNYHLNDWMEERYHIPGVEYVRFYPCSWTFYHHLRQELIOEF 169
DB 121 SLPLVHSDTAKKIPWQNYHLNDWMEERYRHI-PGHFVRFYPCSWTFYHHLRPEILKEF 179
QY 170 TLDHVREREAQKFLRGLOVNGSRPGTFVGVHVRGDDVHVMPKVGWVADRYLQOALD 229
DB 180 TLDHVREREAQKFLRGLOVNGSRPGTFVGVHVRGDDVHVMPKVGWVADRYLQOALD 239
QY 230 WFRARYSSLIFFVTSNGMACRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNHTMTI 289
DB 240 MFRARYSSPVFVTSNGMACRENINASRGDVVFAGDIEGSPAKDFALLTQCNHTMTI 299
QY 290 GTFGIWAAYLTGGDTIYLYANTLPDPSFLKIFKPEAAFLPEWTGIAADLSPLK 344
DB 300 GTFGIWAAYLAGGDTIYLYANTLPDPSFLKIFKPEAAFLPEWVGIPADLSPLK 354

RESULT 12
Q9R275 PRELIMINARY; PRT; 354 AA.
AC Q9R275;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase B.
GN FTB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BDIX;
RA Bureau V., Marionneau S., Cailleteau-Thomas A., Le Mouillac-Vaidye B.,
```

```
RA Liehr T., Le Pendu J.;
RT "Comparison of the three rat GDP-L-fucose: beta-D-galactoside 2-
RT alpha-L-fucosyltransferases FTA, FTB and FTC.";
RL Eur. J. Biochem. 268:1006-1019(2001).
DR EMBL; AF131238; AAD24469.1; -.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 354 AA; 39983 MW; 123E8C8379E8559E CRC64;

Query Match 80.1%; Score 1491; DB 11; Length 354;
Best Local Similarity 77.7%; Pred. No. 9.4e-130;
Matches 276; Conservative 30; Mismatches 37; Indels 12; Gaps 3;

QY 1 MLVQMPFPMAHFLFVFTVSTIFHVQORLAKIQAM--WELPV-----QIPVLA 49
DB 1 MASAQVPFSEPLAHFLFVFTVSTIHLQORIVKLOPLSEKELPMITQSSSGNTESPEMR 60
QY 50 STSKALGPSOLRGWMTINAGRLGNQMGVEYATLYALAKMNGRPAFIPAQMHSILAFIRI 109
DB 61 RDSEQHNGELRGWMTINAGRLGNQMGVEYATLYALAKMNGRPAFIPAQMHSILAFIRI 120
QY 110 TLPVLSATASRIIPWQNYHLNDWMEERYHIPGVEYVRFYPCSWTFYHHLRQELIOEF 169
DB 121 SLPLVHSDTAKKIPWQNYHLNDWMEERYRHI-PGHFVRFYPCSWTFYHHLRPEILKEF 179
QY 170 TLDHVREREAQKFLRGLOVNGSRPGTFVGVHVRGDDVHVMPKVGWVADRYLQOALD 229
DB 180 TLDHVREREAQKFLRGLOVNGSRPGTFVGVHVRGDDVHVMPKVGWVADRYLQOALD 239
QY 230 WFRARYSSLIFFVTSNGMACRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNHTMTI 289
DB 240 MFRARYSSPVFVTSNGMACRENINASRGDVVFAGDIEGSPAKDFALLTQCNHTMTI 299
QY 290 GTFGIWAAYLTGGDTIYLYANTLPDPSFLKIFKPEAAFLPEWTGIAADLSPLK 344
DB 300 GTFGIWAAYLAGGDTIYLYANTLPDPSFLKIFKPEAAFLPEWVGIPADLSPLK 354

RESULT 13
Q9JL27 PRELIMINARY; PRT; 347 AA.
AC Q9JL27;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Alpha(1,2)fucosyltransferase FUT2 (EC 2.4.1.69).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/OLA;
RX MEDLINE=21316545; PubMed=11323419;
RA Domino S.E., Zhang L., Lowe J.B.;
RT "Molecular Cloning, Genomic Mapping, and Expression of Two Secretor
RT Blood Group alpha (1,2)fucosyltransferase Genes Differentially
RT Regulated in Mouse Uterine Epithelium and Gastrointestinal Tract.";
RL J. Biol. Chem. 276:23748-23756(2001).
DR EMBL; AF214656; AAF45146.1; -.
DR MGD; MGI:109374; Fut2.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 347 AA; 39243 MW; DD5F441046CC8E79 CRC64;

Query Match 79.8%; Score 1475.5; DB 11; Length 347;
Best Local Similarity 79.3%; Pred. No. 3e-129;
Matches 276; Conservative 29; Mismatches 38; Indels 5; Gaps 2;

QY 1 MLVQMPFPMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTKALGPSQ- 59
```



```

Db 1 MASAOVPSFPLAHFLIFVFTVSTIIHQRIYVKLQTLSEKELQAVOMSPNAARDIMQ 60
QY 60 ---LRGWTNAINAGRLGNOMGEYATLYALAKMNGRPAPFPAQMHSTLAPFRITLPVLHS 116
Db 61 SAKLOGIFTINSIGRLGNOMGEYATLYALAKMNGRLAFIPESMHNALAPFRISLPVLHS 120
QY 117 ATASRIPWQNYHLNDWMEERYRHPGEGYVRFYTPGCSWTFYHHLRQELLOEFTLHDHVR 176
Db 121 DTARRIPWQNYHLNDWMEERYRHI-PGQYVRFYTPGCSWTFYHHLRPEILKEFTLHDHVR 179
QY 177 EEAQKFLRGLQVNGSRPGTFVGVHVRGDDYVHVPKWKGVADVADRYLOQALDWRFRAYS 236
Db 180 EEAQFLRGLRVNGSQSPSTFVGVHVRGDDYVHVPKWKGVADVADRYLOQALDWRFRAYS 239
QY 237 SLIFVVTNSGMWCARENIDTSHGDVVFAGDIEGSPAKDFALLTQCNHTIMTIGTFGIWA 296
Db 240 SPVFVVTNSGMWCARENINSLGDDVVFAGDIEGSPAKDFALLTQCNHTIMTIGTFGIWA 299
QY 297 AYLTGSDTIYLANITLPSDFPKIKFPAEAFLEPWTGIAADLSPLLK 344
Db 300 AYLAGGDTIYLANITLPSDFPKIKFPAEAFLEPWTGIAADLSPLLK 347

RESULT 14
Q9JK44 PRELIMINARY; PRT; 380 AA.
ID Q9JK44 AC Q9JK44
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Alpha 1-2 fucosyltransferase.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RA Sherwood A.L., Stroud M.R., Levery S.B., Holmes E.H.;
RT "An amino acid region at the N-terminus of rat hepatoma H35 cell GDP-
RT fucose:GMI alpha 1-2 fucosyltransferase is required for optimum enzyme
RT activity and interaction with lipids."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF264005; AAF72200.1;
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11;
DR Glycosyltransferase; Transferase.
KW SEQUENCE 380 AA; 42742 MW; 920692C96689C05C CRC64;

Query Match 79.7%; Score 1475; DB 11; Length 380;
Best Local Similarity 77.7%; Pred. No. 3.8e-129;
Matches 275; Conservative 30; Mismatches 37; Indels 12; Gaps 3;

QY 1 MLVQMPSFPAHILFVFTVSTIIHQRIYVKLQTLSEKELQAVOMSPNAARDIMQ 49
Db 1 MASAOVPSFPLAHFLIFVFTVSTIIHQRIYVKLQTLSEKELQAVOMSPNAARDIMQ 60
QY 50 STSKALGPSQRLGWTNAINAGRLGNOMGEYATLYALAKMNGRPAPFPAQMHSTLAPFR 109
Db 61 RDSQHGNGELRGMTINSIGRLGNOMGEYATLYALAKMNGRLAFIPESMHNALAPFR 120
QY 110 TLPLVHSATASRIPWQNYHLNDWMEERYRHPGEGYVRFYTPGCSWTFYHHLRQELLOE 169
Db 121 SLPLVLSDFAKKIPWQNYHLNDWMEERYRHI-PGHEVRFYTPGCSWTFYHHLRPEILKE 179
QY 170 TLHDHVEEAQKFLRGLQVNGSRPGTFVGVHVRGDDYVHVPKWKGVADVADRYLOQALD 229
Db 180 TLHDHVEEAQKFLRGLRVNGSQSPSTFVGVHVRGDDYVHVPKWKGVADVADRYLOQALD 239
QY 230 WFRARYSSLIFFVVTNSGMWCARENIDTSHGDVVFAGDIEGSPAKDFALLTQCNHTIM 289
Db 240 MFRARYSSPFFVVTNSGMWCARENINASRGDDVVFAGDIEGSPAKDFALLTQCNHTIM 299

```

```

QY 290 GTGCIWAAYLTGGDTIYLANITLPSDFPKIKFPAEAFLEPWTGIAADLSPLLK 343
Db 300 GTGCIWAAYLAGGDTIYLANITLPSDFPKIKFPAEAFLEPWTGIAADLSPLLK 353

RESULT 15
C70504 PRELIMINARY; PRT; 347 AA.
ID C70504 AC C70504
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase (EC
DE 2.4.1.69).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RA Lin B., Hayashi Y., Saito M., Sakakibara Y., Yanagisawa M.,
RA Iwamori M.;
RT "GDP-fucose: beta-galactoside alphaL,2-fucosyltransferase, MFUT-11,
RT and not MFUT-1 or -III, is induced in a restricted region of the
RT digestive tract of germ-free mice by host-microbe interactions and
RT cycloheximide."
EL Biochim. Biophys. Acta 1487:275-285(2000).
DR EMBL; AF064792; AAC16887.1;
DR MGD; MGI:109374; Fut2.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11;
DR Glycosyltransferase; Transferase.
KW SEQUENCE 347 AA; 39215 MW; 616CCB1158104179 CRC64;

Query Match 79.5%; Score 1471.5; DB 11; Length 347;
Best Local Similarity 79.0%; Pred. No. 7e-129;
Matches 275; Conservative 30; Mismatches 38; Indels 5; Gaps 2;

QY 1 MLVQMPSFPAHILFVFTVSTIIHQRIYVKLQTLSEKELQAVOMSPNAARDIMQ 59
Db 1 MASAOVPSFPLAHFLIFVFTVSTIIHQRIYVKLQTLSEKELQAVOMSPNAARDIMQ 60
QY 60 ---LRGWTNAINAGRLGNOMGEYATLYALAKMNGRPAPFPAQMHSTLAPFRITLPVLHS 116
Db 61 SAKLOGIFTINSIGRLGNOMGEYATLYALAKMNGRLAFIPESMHNALAPFRISLPVLHS 120
QY 117 ATASRIPWQNYHLNDWMEERYRHPGEGYVRFYTPGCSWTFYHHLRQELLOEFTLHDHVR 176
Db 121 DTARRIPWQNYHLNDWMEERYRHI-PGQYVRFYTPGCSWTFYHHLRPEILKEFTLHDHVR 179
QY 177 EEAQKFLRGLQVNGSRPGTFVGVHVRGDDYVHVPKWKGVADVADRYLOQALDWRFRAYS 236
Db 180 EEAQFLRGLRVNGSQSPSTFVGVHVRGDDYVHVPKWKGVADVADRYLOQALDWRFRAYS 239
QY 237 SLIFVVTNSGMWCARENIDTSHGDVVFAGDIEGSPAKDFALLTQCNHTIMTIGTFGIWA 296
Db 240 SPVFVVTNSGMWCARENINSLGDDVVFAGDIEGSPAKDFALLTQCNHTIMTIGTFGIWA 299
QY 297 AYLTGSDTIYLANITLPSDFPKIKFPAEAFLEPWTGIAADLSPLLK 344
Db 300 AYLAGGDTIYLANITLPSDFPKIKFPAEAFLEPWTGIAADLSPLLK 347

```

Search completed: May 27, 2003, 15:10:47  
Job time : 37.6774 secs